Phenotypic heterogeneity in high-grade serous ovarian cancer: evidence for a stochastic model of proliferation

Douglas A. Hall

Wolfson College

University of Cambridge

Dissertation submitted for the degree of Doctor of Philosophy

August 2017

Contents

List of Figures

List of Tables

List of abbreviations

(E)GFP	(Enhanced) Green Fluorescent Protein
BrdU	5-Bromo-2′-deoxyuridine
BSA	Bovine Serum Albumin
CAF	Cancer-Associated Fibroblasts
CC3	Cleaved-Caspase-3
CD90	Cluster-of-Differentiation 90 (protein product of the *THY1* gene)
CRUK	Cancer Research UK
CSC	Cancer Stem Cell
DAPI	4′6-Diamidino-2-Phenylindole
DDR	DNA Damage Response (or Repair)
DT	median Division Time
EdU	5-Ethynyl-2′-deoxyuridine
EMT	Epithelial-to-Mesenchymal Transition
EpCAM	Epithelial Cell Adhesion Molecule
FACS	Fluorescence Activated Cell Sorting (flow cytometry sorting)
H+E	Haematoxylin and Eosin staining
HGSOC	High-grade serous ovarian cancer
IHC	Immunohistochemistry
MOI	Mulitplicity Of Infection
MRD	Minimal Residual Disease
NSG	NOD-SCID Gamma (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ) mouse
OSE	Ovarian Surface Epithelium
PARP	Poly-ADP Ribose Polymerase
PBS	Phosphate-Buffered Saline
PDPN	Podoplanin
PDX	Patient-Derived Xenograft
PI	Propidium Iodide
PK	Protein Kinase
QC	Quality Control
RFP	Red Fluorescent Protein
RT	Room Temperature
RT-PCR	Reverse Transcription Polymerase Chain Reaction
scRNAseq	single-cell RNA sequencing
SE	Standard Error
STIC	Serous Tubal Intraepithelial Carcinomas
t-SNE	t-distributed Stochastic Neighbor Embedding
TA	Transit Amplifying cell
TAm-Seq	Tagged-Amplicon Sequencing
TIC	Tumour-Initiating Cell
UMI	Unique Molecular Identifier

Declaration

This dissertation is not substantially the same as any I have submitted, or have concurrently submitted/am submitting, for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or concurrently submitted/am submitting, for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.
This dissertation does not exceed the 60,000 word limit prescribed by the Degree Committee for the School of Clinical Medicine.
This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text and detailed below:

- All analysis-only flow cytometry was my work, but operation of the FACS Aria and INFLUX flow cytometry sorters was performed by members of the CRUK Cambridge Institute Flow Cytometry Core
- Routine histology was performed by the CRUK Cambridge Institute Histology Core as part of their pipeline, and covers all H&E and IHC described here. All immunofluorescence was my own work
- Leigh-Anne McDuffus and Cara Brodie from the CRUK Cambridge Institute Histology Core performed analyses using the AxioScan instrument, and optimised cell-calling using the HALO software
- DNA extracted from patient tumour, and xenograft samples, was used in several projects within the Brenton Laboratory and extractions were therefore performed by various members of the Brenton Laboratory including me
- Tagged-amplicon sequencing and shallow whole genome sequencing was performed as part of other projects by Maria Vias and Ania Piskorz; informatic analysis and formatting of the data in the sWGS validation was provided by Dilrini De Silva
- DNA sequencing, and library preparation for the single-cell mRNA sequencing, was performed by Paul Copeland of the CRUK Cambridge Institute Genomic Core
- Bioinformatics support was provided by the CRUK Cambridge Institute Bioinformatics Core (mostly by Stephane Ballereau); their input to the analysis of the single-cell mRNA sequencing is detailed in the text
- Gopi Shah from the CRUK Cambridge Institute Microscopy Core optimised and operated the Lightsheet Z.1 and performed initial analysis using the Imaris software

Acknowledgements

First and foremost, I would like to thank my supervisors Dr James Brenton and Dr John Stingl, the Stingl and Brenton labs, the CRUK Cambridge Institute as a whole, and CRUK for funding this work. I would like to thank James for all he has taught me: scientific and clinical, but also in terms of data analysis, leadership, and independent project management. I would like to thank the Brenton Laboratory for sharing their repertoires of expertise, protocols, and clinical samples. I would like to thank John and the Stingl Laboratory for all of their guidance, particularly with regard to for instilling good ideals of scientific rigour from the very beginning of my time here, and for sharing their significant technical expertise. I would also like to thank John for his dedication in continuing to provide input and critical appraisal for this project even after leaving the CRUK Cambridge Institute for Vancouver. I would like to thank Siru Virtanen for laying the groundwork within this niche field with her own PhD project, and for exposing an excellent set of questions which this project set out to answer. I would also like to thank Alasdair Russell from the Stingl Laboratory for mentoring me in the early days of my PhD project, and for continuing to offer advice and a sounding-board. Thanks are also due to Ben Simons whose incisive analytical guidance significantly shaped the technical design of the project. I would also like to thank all of the patients on the OV04 study who have contributed to the Lab's ongoing work.

I would like to thank all of my colleagues at the CRUK Cambridge Institute for sharing their extensive expertise with me over the course of my time here, and particular the Core Facilities for creating an environment where research can progress so efficiently. I would like to thank the Flow Cytometry Core, particularly Richard, Nina, and Mateusz for sharing their expertise throughout my time here, and in particular for their guidance (and patience) during my early days of learning the technique. From the Histology Core I would like single out Jodi for sharing her expertise with IHC for application to my IF work, and Leigh-Anne and Cara for their optimisation work of the AxioScan and HALO algorithms, as well as their patience in providing Core support (and moral support) for all the early failed iterations these experiments were built on. I would like to thank the CRUK Cambridge Institute Biological Resources Unit for supporting my *in vivo* work though care of mice, and for their input on technical and Home Office compliance matters. Particular thanks are due to Matt Clayton and Mike Mitchell, to Jeleana Dutton for her help in keeping our paperwork in order, and to Maria Vias, John Stingl and Alasdair Russell for managing the licences under which I have worked. I would like to thank Paul Copeland from the Genomics Core for his expertise and for his work on the library preparation for the single-cell mRNA sequencing, and Stephane Ballereau from the Bioinformatics Core for his work on the analysis and interpretation of the resulting data, along with Rory Stark for his oversight of and input into the latter.

I am grateful to Ann-Sofie Thorsen for sharing her expertise and method development in preparing cells for light-sheet imaging, which allowed me to collect key apoptosis data for this project within its timeframe, and to Gopi Shah for working with the Lightsheet Z.1 and

Imaris software. I am also grateful to Scott Lyons for sharing his LV-indLS2 construct. I would like to thank a range of PIs for short but valuable discussions, either at conferences or by email, particularly Peter Eirew, Rob Clarke, Tariq Enver, Jacqui Shields, and Rebecca Fitzgerald; I would also like to thank Ian Mackenzie for organising the excellent annual UK Cancer Stem Cell Symposium, which I have attended for the last four years, and which is an ideal crucible of ideas for this field. I am also grateful to Brenton Laboratory members Debbie Sanders and Maria Vias for sharing novel 3D-organoid cultures and cell-lines respectively, to Maria Vias and Ania Piskorz for sharing a wealth of data from their work which supported the validity of models I developed during this project, and to Gayathri Chandrasekaran whose ascites work highlighted to me the value of ascites PDX models.

Finally, I would like to thank Wolfson College and the University for supporting me in my studies over the last four years, and the patience of my family and friends.

Abstract

High-grade serous ovarian cancer (HGSOC) represents the largest and most lethal subtype of ovarian cancers. Chemotherapy is effective at depleting these tumours, sometimes eliminating all visible disease, however the tumour invariably regrows from surviving tumour cells. Improving treatment depends on understanding and targeting these residual cells.

The bulk of HGSOC tumour cells strongly express the epithelial surface marker EpCAM. However, previous work in the Brenton and Stingl Laboratories identified a distinct subpopulation of cells with low levels of EpCAM. These cells are uniquely able to regrow tumours in the lab, and are enriched by chemotherapy.

The objective of this project was to understand the dynamics of these cells, not during transplantation or initiation, but under conditions similar to an established tumour in the patient. To this end, tumours were engrafted into immune-deficient mice and allowed to grow. These were then studied only once they had become established.

Nucleoside tracing studies were used to assay proliferation within each of the subpopulations. These were performed both over the short-term (to determine subpopulation proliferation rate), and over the long-term (to determine resulting expansion of the subpopulations over time). The results show that the subpopulation deficient in EpCAM is seen to divide more slowly in all cases.

Both subpopulations divide stochastically, and their rates of proliferation match the expansion of each subpopulation. This indicates that both subpopulations are self-supporting.

Inducible lineage tracing tools were validated to investigate hierarchy and subpopulation interconversion. Unfortunately, the tools generated for this purpose did not prove reliable and solutions were not available within the scope of this project. An alternative approach was also tested, assessing the symmetry of cell division by immunofluorescence, but is not yet optimised.

To provide a deeper understanding of these phenotypes, single-cell-mRNAseq was performed. Sequencing is only half complete, but initial analysis already appear to indicate distinct gene expression patterns between cell clusters corresponding to observed phenotypic subpopulations. This provides a basis for future work into targeting the resistant subpopulation.

In conclusion, this work refutes the proposed cancer stem cell model in HGSOC and presents evidence of stochastic division and self-supporting dynamics in distinct phenotypic subpopulations. It provides evidence of a distinct subpopulation, low in EpCAM, which is slow-cycling. This further supports the idea that these cells are the origin of relapse. In addition, this work begins the process of assessing (at single-cell level) differential expression in this fraction, elucidating the molecular nature of this subpopulation, and guiding future work to target these cells.

Chapter 1

Introduction

1.1 High-grade serous ovarian cancer

Ovarian cancer is the sixth most common cancer type in women, with 7,400 diagnoses per year in the UK, and 4,100 deaths. The five-year-survival rate for all types of ovarian cancer stands at 40% (Vaughan et al., 2011).

Ovarian cancer is a broad term covering an array of different cancer subtypes with very different underlying aetiologies and mechanisms (Gilks, 2004) (Soslow, 2008) (Kurman & Shih, 2010) (Kurman & Shih, 2011). As a result, these subtypes also have their own distinct prevalences and mortality rates (McGuire et al., 2002). The vast majority of ovarian cancers are epithelial, and these are split into 4 subtypes: serous, clear-cell, endometrioid, and mucinous (Gershenson, 2010). Clear-cell and endometrioid tumours are characterised by mutations in *PIK3CA* (Shayesteh et al., 1999) (Campbell et al., 2004) and *ARID1A* (Mao & Shih, 2013) (Jones et al., 2010) (Wiegand et al., 2010) (Chandler et al., 2015), and represent so-called 'type I' tumours (Landen et al., 2008), characterised by less advanced stage at diagnosis, less frequent metastases, and only gradual accumulation of mutations (typically in regulation of growth and metabolism, which directly cause oncogenesis) (McGuire et al., 2002).

Serous disease (representing a little over half of epithelial ovarian cancers) is further subdivided into high-grade and low-grade disease. Low-grade serous represents another localised type I disease, that is typically driven by *BRAF* and *KRAS* mutations. It is also typically detected early, and has comparatively good survival rates if appropriately treated (Kaldawy et al., 2016). However low-grade represents a relatively small proportion of serous cases, and its effective and established treatment options make it a less pressing research priority. High-grade serous disease represents a 'type II' disease (Landen et al., 2008) characterised by aggressive and metastatic spread, late-stage detection, and hypermutation as a result of

genomic instability. Mucinous is rarer, making up less than 10% of cases, and while strictly a type I disease (driven predominantly by *KRAS* mutation) the subtype shares an aggressive clinical nature resulting in progression and mortality more in line with expectations of high-grade serous (Brown & Frumovitz, 2014).

Early work, done before the importance of these subtypes became clear, treated ovarian cancers as a single pathology, rather than a series of distinct diseases unified only by common anatomical location, and thus struggled to find the common themes that would be expected of a single disease. More recent work tackling individual subtypes has begun to make headway, and the field is now heading increasingly in the direction of stratification – further subgrouping within the subtypes themselves, based on common features and drivers in a bid to gain a deeper understanding of their mechanisms, and seeking to identify novel therapies which, while they may not be effective against an entire subgroup, may show significant effects on certain stratifications within it.

High grade serous ovarian cancer (HGSOC) represents the largest subgroup of ovarian cancers, and is responsible for the majority of deaths. Five-year-survival for HGSOC stands at around 30%. The disease is driven by ubiquitous early mutation of *TP53*, leading to dysregulation of genomic maintenance (Ahmed et al., 2010). In around 10% of cases this is supplemented by defects in the *BRCA1/BRCA2* genes which form part of the DNA damage repair (DDR) network.

There are few other common mutational changes in HGSOC. Cyclin E is observed to be upregulated in a subset of patients, and these patients are observed to have worse than average prognosis (Rosen et al., 2006). While the role of cyclin E in the cell cycle gives it obvious oncogenic potential, other mechanisms of action are possible, as Cyclin E has also been implicated in induction of genomic instability (Reed et al., 1999), the key hallmark of HGSOC.

The reason for a lack underlying mutations, is that HGSOC is primarily a disease driven by changes in genomic copy number. The initial mutation in p53 results in cells losing the ability to coordinate proper DNA repair processes. This causes them to gradually accumulate defects, not through point mutations, but through duplication and/or deletion of sizable genomic fragments. This process occurs across the entire genome resulting in wildly fluctuating numbers of copies of any given part of the genomic sequence, as shown in Figure 1.1. These result in erratic changes in the number of copies of a huge range of genes (Bell et al., 2011), with an inevitable knock-on effect for proper translation. This widespread 'genomic chaos' not only feeds back into further genomic instability, but also drives dysregulation of proliferation and ultimately tumour spread and metastasis. Other projects within the Brenton Laboratory are currently investigating patterns of genomic instability ('copy number signatures'), in order to stratify patients based on underlying instability mechanisms (Macintyre et al., 2018). Different signatures have been shown to correlate with different survival rates in patients, which could lead to more accurate prognoses for future HGSOC patients. In addition, by better stratifying the underlying mechanisms driving the disease,

future therapeutics can be better targeted to patients most likely to benefit from their effects, allowing efficacy to be determined more accurately, accelerating the translation of future therapies to clinical practice (discussed further in section 1.7).

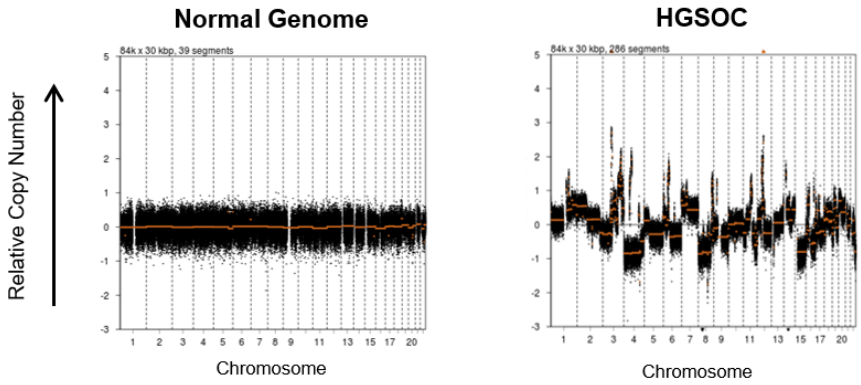


Figure 1.1: Shallow whole-genome sequencing shows the wildly eratic variation in genomic copy number seen in HGSOC. Relative copy number estimates the copy number of a genomic region relative to the average copy number across the whole.

Standard-of-care treatment for HGSOC is carboplatin chemotherapy and possible surgical resection of the disease. Carboplatin (like its predecessor cisplatin) is a DNA-binding/crosslinking agent which disrupts DNA replication, preferentially killing rapidly proliferating cells, and those with DNA repair defects. Several other chemotherapeutics may also be used such as paclitaxel (which disrupts the mitotic spindle preferentially killing cells unable to repair this damage), and gemcitabine (another DNA damaging agent). This chemotherapy is typically effective at reducing the disease bulk, and sometimes eliminates all visible disease. However, the tumour almost invariably regrows from undetectable populations of surviving tumour cells known as the minimal residual disease (MRD), which must either be inherently chemoresistant, or benefit from some extrinsic protection as a result of local microenvironment.

The regrowing disease can be subsequently re-treated with chemotherapy, but chemotherapy resistance typically emerges quickly, after which tumour growth and metastasis cannot be effectively countered. Improving treatment depends on understanding how the MRD subpopulation survives carboplatin therapy, and how this might be prevented. If the factors protecting the MRD cells could be abrogated, or the MRD cells targeted prior to (or concurrently with) carboplatin therapy, then disease recurence should be severely impacted, or potentially even eliminated altogether.

The importance of targeting the MRD subpopulation is underscored by the effective absence of additional therapeutic options for HGSOC beyond combination carboplatin chemother-

apy. The only major exception to this is that over the last few years, the small minority of patients with *BRCA* mutations may also be treated with PARP inhibitors such as Rucaparib and Olaparib (Konecny & Kristeleit, 2016). The PARP enzyme is involved in the repair of single-strand DNA breaks, and its inhibition in cells with BRCA deficiency (involved in the homologous recombination pathway which repairs double-strand DNA breaks) leads to a synthetic lethal effect in tumour cells. While effective, resistance is seen to subsequently develop to this therapy through secondary *BRCA* mutations (Edwards et al., 2008), or by other mechanisms yet to be fully understood (Ang et al., 2013). Options for targeting angiogenesis of the disease are currently under investigation, but beyond these exceptions, carboplatin remains the mainstay of HGSOC therapy.

1.2 Cell of origin

HGSOC tumours are generally localised on the ovarian surface epithelium (OSE), although in many cases they will also have metastasised to other parts of the peritoneum. In the case of later stage disease, cells are often found free-floating (either singly or in clumps) in fluid accumulated in the peritoneum (known as 'ascites'). Given that localised disease is generally observed on the OSE, a layer of mesothelial cells covering the ovaries, historically this tissue was assumed to constitute the cell-of-origin of HGSOC. In support of this hypothesis, this layer is ruptured during ovulation, and therefore undergoes constant damage and wound-healing response. This type of response has been linked with oncogenesis and cancer (Casagrande et al., 1979). In addition, there is evidence to suggest that incidence of HGSOC is correlated with number of ovulation cycles, and that risk is reduced by factors which inhibit this (e.g. pregnancy, and contraception which inhibits the normal ovulatory cycle).

Around 20 years ago, the assumption that the ovary was the tissue of origin of ovarian tumours (not just in regard to HGSOC) began to be challenged. This was primarily based on the morphology of the disease which has more in common with Müllerian tissues than the mesothelial OSE (Dubeau, 1999).

Subsequent work has narrowed down the likely origin to the epithelium of the fallopian tube. The evidence-base for this comes from a clear chain of intermediates between normal cuboidal fallopian tube epithelium and full HGSOC. Normal fallopian tubes sometimes show patches of cells containing a specific *TP53* mutation (the key early driver in HGSOC). The same mutation is seen across the patch (implying clonality). These pre-cancerous features are termed 'p53 signatures'. It is believed further oncogenic mutations result in these signatures developing into serous tubal intraepithelial carcinomas (STICs), and from there into full HGSOC.

There is evidence that p53 signatures give rise to STICs, from both statistical correlation and direct observation. Fallopian tube samples from HGSOC patients are observed to have

higher rates of p53 signatures than those from normal controls (even when BRCA status is controlled for) (Lee et al., 2006). In addition, in a few cases, nascent STICs have been observed directly growing out from p53 signatures; both features showed the same mutational history indicating a common origin (Lee et al., 2006).

STIC lesions are observed in the fallopian tubes of patients with HGSOC, and also to a lesser extent in healthy women, particularly those with predisposition to HGSOC (Medeiros et al., 2006). Like p53 signatures, these STIC lesions are characterised by *TP53* mutation, but demonstrate various degrees of deregulation of proliferation, and shortened telomeres (Chene et al., 2013). In contrast there is minimal evidence of equivalent lesions in the OSE, as might be expected if this was the tissue of origin for HGSOC.

A 2013 paper by the groups of Drapkin and Dinulescu reports mouse models in which *Trp53* (murine *TP53*), *Pten* and either *Brca1* or *Brca2* were inactivated by *Pax8*-driven (i.e. fallopian tube specific) Cre expression (Perets et al., 2013). These models developed STIC lesions followed by tumours of the fallopian tube consistent with HGSOC. The morphology and features of these murine tumours matched well with their human equivalents, and in around half of cases, these tumours then metastasised to the ovary. These tumours also show the raised CA-125 and extensive copy-number variation indicative of HGSOC disease. Furthermore, surgical removal of the ovaries in these models did not prevent development of STIC or HGSOC (although did limit metastatic spread), while removal of the fallopian tubes prevented any STIC or HGSOC formation. However, it is worth noting that in the equivalent models in which only *Trp53* and *Pten* were eliminated, full HGSOC was not seen, with the mice getting only localised endometrial tumours. This latter result is unexpected as *BRCA* mutated tumours represent only a small proportion of HGSOC cases in human patients. This result shows that there is still more to understand about the origin of HGSOC, but it is possible that it arises as a result of the differences in the biology of the mouse and human diseases. This uncertainty underlines the importance of future study moving towards working directly on human material, in order to move away from a dependence on weaker murine models. The observation that metastatic spread is limited by removal of the ovaries is very interesting. As previously discussed, one of the reasons the OSE was considered a likely tissue of origin is because it is damaged by each ovulation, and must undergo repair. Wound-healing processes are known to be oncogenic, and number of ovulation cycles is seen to correlate with rates of HGSOC. While the weight of evidence favours a Müllerian origin for the disease, this does not mean that the constant repair of the OSE does not play a role in disease. Wound-healing responses typically involve secretion of repair factors, specific cell-cell interactions, innate immune response, and tissue remodelling processes. The presence of these interactions could drive further oncogenic change in any early HGSOC cells (or even STICs) in the vicinity. Thus even though the OSE appears not to be the tissue of origin, it may retain an important role in promoting HGSOC. In support of this hypothesis, HGSOC and its precursors are predominantly found towards the distal end of the fallopian tube (nearest the ovary), suggesting that the OSE may act as a catalyst for development of full HGSOC, and its subsequent metastasis.

For a deeper insight into these processes, it is necessary to look to the processes underlying tissue growth and maintenance.

1.3 Cancer stem cells

1.3.1 Tissue stem cells in normal tissue

The generation of a complex multicellular organism is dependent on totipotent stem cells which, throughout development, generate the many diffferent cell lineages found in the mammalian organism. Over the development process, many of these stem cells begin to gradually differentiate and take on roles in the development of specific organs. These cells lose totipotency becoming merely pluripotent and restricted to the lineages of that organ alone. Beyond early development, the body also contains cells with very defined lineage commitment, but which retain the ability to divide indefinitely. These cells, which are limited to only one (or at most a small tree of) lineages, are termed 'tissue stem cells'. These cells are often seen in epithelial tissues, as constant ablation of cells at the epithelial surface leads to a continuous need for replacement cells. Epithelial stem-cells are typically are found in the basal layer of the epithelium, and generate progeny which migrate through the tissue to eventually replace the surface epithelial layer. In many cases, tissue stem cells do not undertake the bulk of the division in the tissue themselves, instead dividing only slowly, but giving rise to a stream of rapidly cycling progenitor (transit-amplifying or TA) cells which undertake the bulk of the division, but which have limited proliferative potential (entering senescence within small number of divisions). This arrangement helps to limit the accumulation of deleterious mutations and genetic damage (associated with division) within the tissue stem cells. The tissue stem cells act as a 'master copy', cycling only slowly, and periodically releasing fast-cycling progenitors to do the bulk of the proliferative work. It is unimportant if these progenitor cells accumulate genetic damage since they will produce only a small batch of cells before entering senescence, and they and their progeny will be lost within a relatively short period, in favour of new cells derived from the tissue stem cell. In addition to their slow rate of proliferation, tissue stem cells appear to activate a number of additional pathways to protect them from genetic damage and to ensure the fidelity of their genome.

The presence of these adult stem cells has been demonstrated in many normal tissues. A range of these are discussed below at some length in order to both demonstrate evidence for the concept, and to showcase the range of techniques available for interrogating this type of biology (as well as to highlight briefly some of the limitations of these methods).

Among the best models of tumour stem cells is the lining of the intestine. The laboratory of Hans Clevers identified a rare population of cells in this tissue, isolated by the protein Lgr5, which uniquely gave rise to all of the epithelial lineages seen in the gut (Barker et al.,

2007). To demonstrate the 'stemness' of these cells, the authors created a mouse expressing an inducible mark in Lgr5$^+$ cells. When induced, this mark was seen in the proposed tissue stem cells. This mark remains post-induction, and was seen to 'spread' away from these cells up the walls of the intestinal crypt in mice culled at longer time periods post-induction. This indicates that these cells lining the walls of the crypt arise from the tissue stem cells at the base, and that these new progeny displace the existing cells of the crypt wall, driving their movement up and out of the crypt. Marking of the crypts using nucleoside analogues allowed the authors to demonstrate that the majority of these cells are undergoing cell cycle within a 24 h period, laying a strong foundation for future studies into the dynamics of intestinal crypts. In addition, these studies also quickly label a band of cells a little way from the bottom of the crypt, implying that this band represents faster-dividing transit-amplifying cells. The same group confirm this in a subsequent paper where they show that knock-out of APC (adenomatous polyposis coli) within the stem cell compartment leads to adenoma formation, while in the TA compartment it causes short-term adenoma-like phenotypes, but these transformations show negligible expansion in the long-term (Barker et al., 2009). This is indicative of transformed TAs – rapid expansion to form limited tumour-like features, followed by an inability to sustain long-term proliferation and lapse of the cell population into quiescence.

Subsequent studies have used lineage-tracing to trace the dynamics of these tissue stem cells, and the resulting data have allowed revised estimates of the number of stem cells per crypt at 5-7 (Kozar et al., 2013). This revision implies that earlier estimates by Lgr5-positivity alone were somewhat inaccurate, suggesting that it may not be fully specific to these tissue stem cells. The ebb-and-flow in clonal dynamics as different clonal 'families' of cells grow and shrink demonstrated that stem cell loss and replacement are relatively slow processes. In combination with nucleoside tracing data, this suggests that the vast majority of stem cell divisions are asymmetric, giving rise to non-stem progeny, and that replacement of tissue stem cells is a relatively uncommon event. A more detailed study of this process shows that accumulation of various oncogenic mutations can lead to favourable survival chances for clones, but that this bias is relatively small, and that thus most mutant clones will still be lost as a result of neutral drift (Vermeulen et al., 2013). Furthermore, this study demonstrates that the crypt-based structure of the intestinal epithelium inhibits takeover by mutant clones, as once a clone has dominated the base of a crypt, it is much harder for it to expand further (a process which occurs primarily by crypt fission, and competition at the level of crypts themselves).

There also appears to be a strong degree of plasticity within apparently strongly defined hierarchies: investigation of the very slow-cycling so-called '+4' cells (which are Lgr5$^+$, but do not seem to be the same as the tissue stem cells) indicate they are also stem-like but restricted to a distinct lineage of cells called Paneth cells found at the base of intestinal crypts (Buczacki et al., 2013). This may explain why assessment of tissue stem cells by Lrg5-positivity overestimates the number of tissue stem cells. However, upon gut injury (and associated wound healing responses), they appear to lose this lineage restriction, gaining

multipotence and giving rise to the progenitors and differentiated cells of the crypt wall. This plasticity in the face of injury and repair is of particular interest with reference to HGSOC as these conditions replicate the wound-healing processes seen in ovulation, experimental tumour transplantation, and potentially the effects of chemotheraputic toxicity.

The presence of tissue-stem-cells has also been demonstrated in normal (murine) skin (Clayton et al., 2007). This study used sparse-labelling (by titration of a very low concentration of inducing agent into a germline modified mouse) to mark 1 in 600 basal cells of the intrafollicular epidermis. The resulting distribution of clones was then followed 'over time' (i.e. by performing the same experiment in parallel in different mice and culling at different periods post-induction). The number of marked clones was seen to decline stochastically over time, while the number of cells per clone was seen to increase over time (forming a Poisson distribution around an increasing mode). These results are very informative, as a slow and smooth increase in the number of cells arising from each clonal patches is inconsistent with a transit-amplified model. Indeed, the smooth Poisson curve is indicative of a purely stochastic pattern of division from the tissue stem cells, indicating a single dividing pool. The smooth and stochastic drop-off in the number of clonal patches is indicative of neutral-drift, in which clonal patches are lost as their rate of replacement is overcome by the rate of cell attrition, and surrounding clones expand in response. Since each clonal patch is maintained by a stem cell, the loss of clonal patches and epansion of others indicates one stem cell being replaced by another. In order to maintain larger patches, stem cells must give rise to daughter stem cells in order to maintain the increasingly large patch sizes, and in turn these daughters would begin to compete with their parent patches in any subsequent clonal labelling experiment. As apoptosis is demonstrated to be negligible in this experiment, this indicates that tissue stem cells must be able to leave this proliferative state and enter quiescence (in order to explain loss from this pool). To demonstrate this directly, the authors used immunofluorescence imaging to observe cells undergoing division. This experiment shows clear asymmetric divisions, in which proliferation factors such as Ki-67 are localised into only one of the daughter cells. They also show both types of symmetric division, giving rise to pairs of proliferating cells, and pairs of non-proliferating cells. These results show evidence of the existence of single-proliferating-compartment models, and demonstrate the extent to which stochastic regimes can dominate the mathematics of such proliferative processes. This work also shows that rigorous mathematical analysis of a dataset can provide significant insight into the processes underlying tissue maintenance.

Studies of the normal mammary gland have produced conflicting conclusions. Jane Visvader's laboratory report that lineage labelling of basal cells gives rise to 20% labelling of luminal cells over an 8 week period, and report mixed clones in established glands, indicative of bipotency, with basal cells able to give rise to basal and luminal lineage (Rios et al., 2014). However, there has been some criticism of this paper for inducing fluorescent marking at high density, leading some to question whether these multilineage clonal patches are defintiely clonal, or whether they could arise by chance marking of two progenitors (of different lineages) in close proximity, which then expand and merge. These studies also

demonstrate evidence for the presence of luminal transit-amplifying cells, which generate small but transient clonal patches which over time are replaced by unmarked cells.

In contrast, other lineage tracing studies suggest that the gland is maintained by distinct basal and luminal cells, as induction of fluorescence within one compartment does not result in labelling in the other (Van Keymeulen et al., 2011). Unlike the previous report, this study presents evidence to confirm that labelling of basal cells through selective promoters is actually specific to this lineage, which appears to be another shortcoming of the previous study, which did not present controls to demonstrate that immediately after labelling there was no off-target marking of luminal populations. Of particular interest to this project, the authors also looked at whether this observation held in implantation/gland initiation studies, to determine if some contradictory data may have been a result of limitations in this method. They implanted mixed luminal and basal cell populations after labelling each in separate experiments, and observed that each labelled population only gave rise to its own lineages. However, when marked basal cells were implanted in isolation, they resulted in normal glands in which both lineages carried the mark specific to the basal cells (luminal cells in isolation do not form glands). This result demonstrates that while lineage specific proliferation may be the norm under established physiological conditions, some cells appear capable of an increased range of proliferation options when removed from those ideal conditions. This is an important observation which will guide the design of this project.

The limits of transplantation/initiation studies are also highlighted by more detailed assessment of progenitor pools within subcompartments of the luminal population (Giraddi et al., 2015). Within the luminal fraction is a subset of 'non-clonogenic luminal cells', named for the fact they do not show any proliferation in transplantation studies. The authors of this paper show, by short- and long-term nucleoside tracing studies, that this population is actually proliferating rapidly when studied *in situ*, and that its inability to initiate clonal expansion post-transplantation appears to simply be an artefact of the nature of this assay.

Assuming the extensive literature in favour of a fallopian origin for HGSOC is accurate, it would also be of value to understand the dynamics of tissue regeneration in normal fallopian tissue. A 2012 paper by Sanaz Memarzadeh's laboratory reports enrichment of tissue stem cells (based on spheroid-forming capacity) in cell fractions deficient in TUBB4, β-tubulin, and PAX8 (Paik et al., 2012). The fallopian tube is composed of two main cell types: ciliated (marked by TUBB4 and β-tubulin), and secretory (marked by PAX8). The authors suggest that a third cell type, the intercalary (or 'peg') cell subset, which does not express either of the above markers, represents some form of tissue stem cell, and may be the origin of HGSOC. These cells were isolated by their expression of EpCAM (confirming they are epithelial cells), along with CD44 and integrin alpha-6 (ITGA6) which appear unique to this population. This population resulted in 5x as many spheroids *in vitro*, suggesting that this population is enriched for cells capable of clonal expansion, but that they are not unique to it (within the limits of this assay). The authors also provide additional circumstantial evidence for some form of involvement, in that this cell type is also reported to be enriched

towards the distal end of the fallopian tube.

While it does not appear to be the cell of origin, the OSE is also of interest given the typical localisation of localised HGSOC disease. It has been known for some time that the density of Müllerian-like transformations on the ovarian surface increases towards the hilum – the region of the OSE where it connects to the fallopian tube (Dubeau, 1999). In 2013 the Nikitin Laboratory presented evidence from mouse models that the hilum (the junction between the fallopian tube and the surface epithelium of the ovary) may be the initiating niche of HGSOC (Flesken-Nikitin et al., 2013). Within this region, the authors identified slow-cycling cells (positive for the stem-cell markers Aldh1, Lgr5, and Cd133) which appear much more active in both *in vitro* and *in vivo* proliferation assays. Lineage tracing reveals that the cells of the hilum give rise to the rest of the ovarian surface epithelium, and mutation of key oncogenes (*Trp53* and *Rb1*) within the hilum results in oncogenesis, while similar mutations in the rest of the OSE are not observed to have this effect. Activation of proliferation is shown to be in response to oestrus, providing a rationale both for the presence of these tissue-stem-cells (repair post-ovulation), and for why oncogenesis might commonly occur here (repeated wound-healing responses and a pro-proliferative niche).

This demonstrates strong evidence for the presence of progenitor cells within the hilum niche of the OSE, and shows that they are apparently uniquely responsible for the repopulation of the entire OSE. This paper does not present any evidence directly connecting these cells to the initiation of HGSOC, but the circumstantial spatial (hilum and distal end of the fallopian tube) and temporal (link to ovulation and associated HGSOC risk) correlation suggests that there is a very real possibility that these cyclic bursts of proliferation may act as an enabling niche for HGSOC. This could occur as a result of either HGSOC or pro-HGSOC (STIC-like) cells being shed from the fallopian tube and embedding in the OSE stem cell niche in the hilum, the diffusion of pro-proliferation repair factors into the distal end of the fallopian tube, or possibly even damage to the fimbae (the extensions of the tube which surround the ovary) as a result of the rupture of the OSE. If Paik *et al.* are correct and the fallopian tube does indeed contain its own tissue stem cells, then these would be better placed than most to be the origin of HGSOC

1.3.2 Cancer stem cells

The tissue stem cell theory raises several important questions as to the possible role of these cells in cancer, and has resulted in two important hypotheses. Tissue stem cells are characterised by a number of features, in particular their indefinite replicative potential, and resistance to external senesence and apoptosis factors. These traits are several of the key hallmarks of cancer (Hanahan & Weinberg, 2000). The 'cell of origin hypothesis' conjectures that these tissue stem cells are effectively several steps down the road of oncogenesis, and should have less mutational distance to cover in order to become oncogenic. As such it proposes that many cancers may result from oncogenesis of tissue stem cells, as they gain

the additional requisite features including accelerated division, rather than arising from non-dividing cells which have overcome their natural barriers to unlimited division. In addition, if untransformed tissue is repopulated via a hierarchical tissue stem cell model of proliferation, it is entirely possible that such proliferative organisation persists into the tumour itself. This is the 'cancer stem cell (CSC) hypothesis', which suggests that tumour consists of transformed cancer stem cells, which not only self-renew, but proliferate to give rise to other 'bulk' tumour cells which are themselves quiescent. It can be argued that if the cancer stem cell hypothesis is true, then it provides indirect support to the cell of origin hypothesis, as it would be surprising if a cancer stem cell hierarchy were recapitulated from oncogenic transformation of a non-tissue-stem-cell, while it is intuitive that a transformed tissue stem cell might generate a warped version of its original hierarchical structure upon transformation. The cancer stem cell hypothesis also raises some questions around the definition of a cancer cell, as a quiescent tumour bulk would not necessarily conform to many of a classical definitions of cancer, for which proliferation is generally considered a prerequisite, suggesting that only the cancer stem cells might truly qualify as cancer cells. One possible piece of circumstantial evidence in favour of the cell of origin hypothesis is the degree of relapse observed in tumours such as HGSOC. If this hypothesis is true, then cancer cells might well be expected to have the properties observed in many tissue stem cells. In particular these might well be chemoresistant. If the cancer stem cell hypothesis is also true, bulk tumour cells might not be chemoresistant, given that this feature may well be arising from hierarchical phenotypic state (i.e. 'stemness'), and not from the oncogenic transformation process. Thus cytotoxic chemotherapy treatment would be expected to kill the bulk of the tumour, but leave the CSCs intact, leaving resistant and proliferative cells from which the tumour would be able to regrow. Tissue stem cells are typically slow-cycling, and this property is one feature which might explain why CSCs would be less affected by platinum chemotherapy (which induces terminal replication stress in cells undergoing rapid division). Several other features of tissue stem cells might also be important. Many tissue stem cells are seen to activate or upregulate pathways relating to drug eflux or DNA repair and maintenance (Moitra, 2015). This makes intuitive sense with respect to the idea that these cells act as high-fidelity genetic blueprints for rapidly dividing tissue. Were these cells to become oncogenic as the cell of origin hypothesis advances, then the same features would also protect them from the effect of platinum chemotherapy, resulting in resistance in small CSC subpopulations, as appears to be observed. As successive rounds of chemotherapy deplete the non-proliferating bulk tumour, the proportion of chemoresistant proliferative cells rises, and susceptibility to chemoresistance falls, mirroring clinical observations around HGSOC relapse. Just as tissue stem cells have been isolated in normal tissues, investigations have been conducted into CSCs using many of the same techniques.

A 2012 study on papillomas, used similar lineage-tracing methods to those used by Clayton *et al.* in the study of normal skin epithelium (discussed previously), but applied to mouse skin subject to carcinogen-induced oncogenesis. This study found similar stochastic patterns underlying papilloma growth, but found a two-population model (stemlike cells and transit-amplifying progenitors) to be a better fit for the data (Driessens et al., 2012). It is unclear

if this is a feature of oncogenesis (i.e. the stemlike cells are not present in normal skin and are simply progenitors acting aberrantly as a result of oncogenic dysregulation) or that oncogenesis amplifies the effects of an existing set of tissue stem cells that were not detectable in the original study. Alternatively, it may simply be that improvements in methodology in the 5 years between the two studies have resulted in improved detection quality.

The laboratory of Luis Parada has demonstrated that markers which identify slow-cycling adult neural stem cells, also identify a subset of glioblastoma cells (Chen et al., 2012). While glioblastoma might seem to have little in common with a disease like HGSOC, the parallels are striking. The disease is aggressive, hypermutational, and p53 mutation is a key driver in the majority of patients (Kyritsis et al., 1995). Treatment is by antiproliferative chemotherapy, which is initially effective but rapidly overcome by resistance. In this paper Chen *et al.* identify a subset of cells which appear to escape anti-proliferative chemotherapy and regenerate the disease post treatment. They also present evidence from tracing experiments that appear to show that these cells represent a cancer stem cell subpopulation able to give rise to the other lineages after ablation of the bulk of tumour cells. Additionally, they demonstrate that selective ablation of the proposed CSC fraction (using an inducible susceptibility gene under the promoter of their proposed CSC marker), leads to elimination of tumours, and significantly improved survival. Together these pieces of evidence provide a very robust case for CSCs within this tumour type, providing a strong foundation for the possibility of CSCs in similar tumours, including HGSOC.

1.3.3 CSCs in HGSOC

Given the challenges of relapse in HGSOC, a number of groups have hypothesised a role for CSCs within the disease. However, at least until very recently, the evidence base in support of a CSC theory for HGSOC has been limited.

An early effort to purify CSCs in HGSOC reported that CSCs could be effectively enriched by selection of cells double-positive for the markers CD44 and CD117 (Zhang et al., 2008). This study looked at serous ovarian cancer cells (although high-grade status for these samples does not appear to have been validated) and their ability to form spheroids in non-adherent culture. The results suggested tumour initiating cells (TICs) were rare, and estimated this rate at 1% (given assumptions of perfect TIC readout in this experiment, that all TICs can sustain spheroids long-term in *in vitro* culture, and that culture conditions that favour tumour-initiation do not promote stemness in non-CSCs). This study also reported upregulation of classical embryonic 'stemness' markers (Oct-4, Nanog, nestin) in spheroids, and that cells could be made to 'differentiate' in response to altered culture conditions. The latter finding raises several questions. Firstly, if TICs represent CSCs, then why do their non-CSC progeny (the bulk of each spheroid) also express stemness markers, since these cells should be non-stemlike. Secondly, if the cells can be shifted from one state to another in response to media changes, then this indicates either a degree of inherent plasticity that

is questionably consistent with a CSC model. It also indicates that cell state is being defined by the culture conditions, suggesting that the results are essentially an artefact of the experimental system.

Another study showed similar results with markers CD44 and MyD88, with this subpopulation being enriched by non-adherent culture, and tumour initiating in immune-deficient mice (Alvero et al., 2009). This study also reported upregulation of pro-inflammatory factors such as IL-8 in these TICs, as well as positivity for CK18. Finally they also observed less sensitivity to carboplatin and paclitaxel in cells populations selected for CD44 positivity.

A handful of other markers have also been implicated as isolating TICs, including aggregation of p53 (Yang-Hartwich et al., 2014), and the so-called 'side-population' (Szotek et al., 2006) which was shown to be colony-forming and more chemoresistant *in vitro*, and induce quicker tumour-initiation in nude mice. A paper a few years later from the same laboratory suggested $CD24^+/CD44^+/EpCAM^+$ cells were TICs (Wei et al., 2010). This is based, among other experiments, on colony-forming assays – while these do show meaningful enrichment of triple-positive cells, this is assayed against triple negative only, which seems a very artificial choice of baseline (rather than just the bulk tumour). In addition, they report this population is enriched by cisplatin or doxorubicin treatment, but decreased by growth inhibition with Müllerian inhibiting substance, however this is again controlled against a triple-negative baseline.

Curley *et al.* reported in 2009 that CD133 enriched for TICs in a handful of patient samples, and demonstrated tumour-initiation from one to two orders fewer cells (Curley et al., 2009). However, there is no additional experimental evidence to support the conclusion that $CD133^+$ cells are Cancer Stem Cells beyond having greater TIC capacity. Kryczek *et al.* replicated a similar degree of enrichment using CD133, but built on this demonstrating that $ALDH^+$ also enriches for TIC capacity, and that cells double positive for these markers have highly enriched TIC capacity (Kryczek et al., 2011). Silva *et al.* report very similar findings, but also show that $ALDH^+$ cells are also enriched by chemotherapy (Silva et al., 2011). In the process, the authors implicitly highlight the limitation of this work – namely that the subpopulation that is enriched by a stressful proliferation assay turns out to be resilient to stress, suggesting that the original assay may not be reading-out the feature it is assumed to be reading out. They also go on to show that enrichment of these markers is associated with poor prognosis in patients, but again enrichment of chemoresistant cells would be expected to be associated with poor prognosis, quite independent of the *in situ* proliferation potential of the cell subpopulations.

In 2011, Stewart *et al.* reported an assessment of the tumour-initiating capacity of HGSOC cells (Stewart et al., 2011). The authors report variability in rates of TICs across a range of patient samples and xenografts. The utility of the marker CD133 in isolating CSCs was tested by flow sorting positive and negative subpopulations, and xenotransplantating each at limiting dilutions. These studies found that CD133 did not reliably enrich for TICs, instead finding that TICs were apparently enriched in different subpopulations at varying

rates, depending on the specific patient sample. The markers CD44, CD117, EpCAM and ALDH1 were also tested and reported to be even less useful. This paper also raises questions about the reliability of xenografts as a model, noting some apparent differences in marker expression between xenografts and original patient tumours.

This paper, along with unpublished work from the Brenton Laboratory and a 2015 paper from the Memarzadeh Laboratory (which will be discussed shortly), represent the latest work investigating CSCs in the context of HGSOC. However, this paper is built on the underlying assumption that a non-CSC model would exclude phenotypic complexity, and that tumour-initiating capacity is a reliable assay for identification of CSCs. While there is certainly evidence to support cases where these two assumptions hold true, this does not mean the same must necessarily be true of all circumstances, and specificaly HGSOC. In addition, it is worth noting that purified subpopulations were xenotransplanted in isolation, so tumour-initiating capacity was being measured in the absence of supporting cells, which is very different from the growth of tumours in the patient when these cells would be surrounded by cells of other lineages which may have important cell-cell interactions. This project sets out an alternative model which embraces the concept of multiple phenotypic subpopulations with differing tumour-initiating capacities, but in which division is fundamentally stochastic and non-hierarchical. This option does not appear to be considered in the papers discussed above.

1.3.4 Mouse models of HGSOC

A number of HGSOC model mice have been developed through targeted mutation of the germline. As previously discussed in section 1.2, these models have proved invaluable in showing that the fallopian tube represents the tissue of origin for HGSOC, as germline induction of driver mutations here create malignancies which reflect the development of HGSOC (Perets et al., 2013), while modifications to the OSE recapitulate less aggressive tumours which do not show the characteristics of HGSOC (Flesken-Nikitin et al., 2013).

However, the further use of these models for translational investigation into human disease is inherently limited by their mouse origins. Fallopian-specific germline mutation models create a range of features that are sufficiently indicative of HGSOC that they strongly support a fallopian tissue of origin. There are however a number of discrepancies between the tumours recapitulated by these mouse models, and what is observed in human disease, and some have suggested that germline mouse models of disease may be less representative than patient-derived HGSOC cell-lines cultured *in vitro* (Morton & Houghton, 2007). For example, Perets *et al.* found that ovarian metastasis required mutation of either *Brca1* or *Brca2*, and that the absence of these mutations prevented spread from the origin site in the fallopian tube to the ovary. This does not mirror human disease, where 90% of HGSOC patients do not have BRCA mutations despite ovarian metastasis being ubiquitous. There are at least two readily apparent potential explanations for this. The first is that mice

have key anatomical differences in the hilum, where the fallopian tube meets the ovary, and these could result in differential rates of metastasis to the ovary. Alternatively, these mouse models survive for only a short window of time (< 1 year), compared to the many years it apparently takes human disease to develop into full HGSOC, and it may be that the mouse models would develop full HGSOC if they existed over time-frames equivalent to the development of the human pathology.

Both the differences in biology, and the time limitations on mutational development imposed by the mouse model, have led to the need for more complex HGSOC models to further investigate the disease. The main limitations on the use of human tissue are ethics and practicality. Human tumour material can be obtained comparatively easily as it is a discarded by-product of therapeutic debulking surgery, but its *in vivo* characteristics can only be studied in an observational manner, as there is no opportunity to perturb the disease *in vivo* and study the response in a controlled fashion. Once *ex vivo*, HGSOC tumour is prone to death, and requires complex conditions to grow at all, quite aside from growing in a fashion that is representative of the original *in situ* disease. While *in vitro* models of surgically recovered human disease exist, the best models of the disease will be ones which best recapitulate the *in situ* conditions in the patient. As such, patient-derived xenografts (PDXs), in which human patient material is surgically implanted into immune-depleted host animals (typically mice) are of significant interest in further investigation of HGSOC.

PDX models can be derived from human patient material in either the form of a piece of solid tumour, or dissociated single cells. The former offers the potential to retain some of the structural complexity, stromal support processes, and cell-cell interactions around the tumour during the implantation process, potentially resulting in better rates of successful implantation, and a better recapitulation of the original tumour. However, it also requires surgical implantation into each mouse which represents a logistical and ethical trade-off. Dissociated cells may be implanted more easily through injection into the host mouse, but with a potential reduction in engraftment efficiency and fidelity. Alternatively, free-floating tumour ascites cells may be collected (as part of therapeutic ascitic drains of a patient's peritoneum) and implanted in the same fashion as single cells. These cells appear capable of surviving in isolation from the tumour, and so are expected to have a higher engraftment rate. However, this comes with a risk that such a model only represents the behaviour of these atypical cells which are presumed to be an advanced form of disease which has potentially undergone metastatic transformation.

The second major choice in the development of a PDX model is whether to implant the tumour in a broadly anatomically equivalent site (an orthotopic xenograft), or at an alternative (heterotopic) site. For HGSOC the former is typically the peritoneum through either intraperitoneal (IP) injection or surgical implantation, although ovarian or fallopian implantation is also possible, while heterotopic implantations tend to use subcutaneous implantation (surgical or injection) on the flanks. The main rationale for the orthotopic approach is to maintain an improved recapitulation of the original human disease site, and

give the tumour access to a similar milieu of signalling factors and hormones. This comes with the trade-off of more complex implantation processes, a higher risk of errors in implantation, and greater challenge in tumour monitoring. The subcutaneous approach allows much better monitoring of tumours, reducing ethical burden, while the significantly reduced logistical burden allows for more complex and extensive experiments with limited resources. The advantages of orthotopic implantation are reduced in the case of HGSOC by several factors. Firstly, the anatomical differences between mouse and human hilum provide a hard limit on the degree of orthotopic equivalence that can ever be achieved in a mouse model. Secondly, most orthotopic models actually use the peritoneum (due to the complexity of implanting directly into fallopian tube or ovary) and as such these models are only really orthotopic for late stage metastatic disease which is inherently already largely site-agnostic. Thirdly, one of the main rationales for orthotopic implantation is to preserve hormone signalling to the disease. However, this rationale is not only limited by the points raised above, and that it assumes mouse hormones are equivalent, but also by recent evidence that hormone signalling has a very limited impact on HGSOC (Sieh et al., 2013). In light of the current cell-of-origin information on HGSOC, this is perhaps not as surprising as it was when the OSE was the suspected site of origin, as the fallopian tube is not subject to the same hormone-driven rapid growth cycles that occur in the OSE. Finally, part of the reason orthotopic implantations have predominated historically has been the improved efficiency of engraftment observed (Hidalgo et al., 2014), which is important in situations where implantation rates impact the viability or reliability of an experiment. However as methods and technology improve, the importance of this consideration has reduced, as improved rates are seen from heterotopic implantations

The use of primary human ovarian xenografts goes back to at least 1977 when Davy *et al.* reported successful serial passage of subcutaneously implanted ovarian material into nude mice (Davy et al., 1977). However, this work was done long before ovarian cancer subtyping was known or accepted, and as such the nature of the material used is questionable.

Work by Weroha *et al.* at the Mayo Clinic has created a large bank of ovarian PDX models using mechanically minced tumour material injected into SCID mice (Weroha et al., 2014). The authors report 168 models created with a 74% engraftment rate. This study is interesting in that the large numbers involved allow for statistically meaningful confirmation of the broadly assumed hypothesis that more advanced disease, particularly that showing potentially metastatic ascites cells, is more likely to engraft. The study found that high-grade disease is much more likely to engraft, as is disease of advanced stage, or that with ascites. As the full dataset is not available it is unfortunately not possible to stratify the set down to confirm these findings in HGSOC alone, but the high prevalence of HGSOC means that this is likely to be the case, particularly for the finding on ascites which is significant to $p < 10^{-3}$ Validation of examples of these PDXs is shown through both immunohistological and genomic assays, and many appear consistent with high-grade, but without appropriate stratification and correlation of these data it is hard to accurately assess reliability. Additionally, while it is clear that stratification has been performed, a number of comments

and references in the published results appear to treat ovarian cancer and HGSOC as interchangeable, which raises a number of concerns with regard to how reliably these models can be used for studies on HGSOC without additional validation. It is also worth noting that these models were grown in SCID mouse models which are less sophisticated than their successor varieties, NOD-SCID, and NSG mice, and may not recapitulate an optimum model.

The laboratory of Ben Neel has reported testing a range of orthotopic and heterotopic methods in the development of HGSOC PDXs, through dissociation and injection into the peritoneal cavity, renal capsule, ovarian bursa, and mammary fat pad of NOD/SCID mice. They report high take rates, averaging 89%, and above 70% for all sites. Furthermore, they also report that successful PDXs were always seen to be successful upon serial passage. As a result of these findings they have moved to using the mammary fat pad model due to ease of implantation, and improved accuracy for detection of small tumours (Stewart et al., 2011).

Curley *et al.* used similar methods, dissociating and injecting into the peritoneum and subcutaneously. Curiously, they report that the intraperitoneal injections resulted in poor success rates and poor tumour development even where it was observed, while their subcutaneous injections gave rise to tumours efficiently (85% success rate for initial engraftment). This runs against conventional wisdom, which suggests that orthotopic sites should produce better results, and may represent another indication that the peritoneum, although typically treated as an 'orthotopic site' may actually be a poor recapitulation of appropriate conditions (Curley et al., 2009).

Kryczek *et al.* also used heterotopic dissociation and subcutaneous injection, reporting generally efficient tumour formation (although they do not report a figure for efficiency). The experimental implantation work described was done directly from human material rather than serial passage PDXs, and at low replicate number, so it is hard to distinguish implantation efficiency from experimental conditions, but it can be extrapolated to be broadly in the region of 80% (Kryczek et al., 2011).

Since the work described in this thesis, a number of additional HGSOC PDX models have been described by various groups, and a few examples are noted here for context.

Liu *et al.* developed PDX models in NSG mice using ascites cells which were implanted in an intraperitoneal fashion (Liu et al., 2016). Tracking these cells *in vivo* required insertion of a luciferase reporter, which in turn required 5-7 days of *in vitro* culturing, although it appears this process was conducted between passages after initial *in vivo* implantation had been completed successfully. The authors report a success rate of 31% for three successive serial passages, and a latency of 2-12 months between engraftment and clinically detectable tumour; they also demonstrate that the resulting PDX models replicate the genomic, immunohistological, and chemotherapeutic response characteristics of the original tumour.

In contrast George *et al.* developed an orthotopic model in NSGs using matrigel-supplemented

surgical implantation (George et al., 2017). They report a 93% engraftment rate, and did not label their cells, assessing by palpation, ultrasound, and PET scanning. Again, no indication of efficiency at the technical level is given, and latency periods are unclear.

There has also been further work published by the Neel Laboratory. The latest publication continues the approach previously discussed, of dissociation and heterotopic injection into the mammary fat pad, but has evolved to take advantage of the latest NSG mouse models. Implantation success rates from the are reported to exceed 90% (Cybulska et al., 2018).

All of these approaches have various merits and drawbacks, primarily a continuum between models that are easier to develop and measure (heterotopic and injection approaches), and those which are more challenging but which may represent a better model (implantation and orthotopic approaches). It is not unexpected that different groups have taken different approaches, because the scientific questions, and practical circumstances will vary in each case. Selection of an appropriate model will always be dependent on the details of a specific study (which potential variables are most likely to impact the hypotheses in question, and which can be more reliably neglected), and the resources available (for example it may be better to develop a number of PDX models at the cost of a little fidelity than generate an ideal model but lack the manpower and resources to produce an adequate group of biological replicates to validate any findings).

1.3.5 The roles of proliferation and apoptois

Proliferation and apoptosis will be discussed in some detail as part of this work, and so some previous literature in the field is presented here for the purpose of providing some context. Both of these processes are naturally of considerable interest across the wider cancer field, as proliferation and apoptosis both offer potential areas of therapeutic leverage, through the inhibition of the former to prevent tumour growth, and release of the latter to kill cancer cells through synthetic lethal effects.

In the HGSOC field, disease-specific research is again limited by recent discoveries of the disease's stratification into subtypes, and that the tissue of origin has been previously misunderstood. Compounding this is that for the purposes of this work, our interest is in quantitative assessment of the rates underlying these processes. However, almost all work on the subject has been directed at the upregulation or downregulation of these molecular pathways (relative to control samples), as is logical given the translational ambition of such studies, rather than any absolute quantitative output.

Proliferation

The proliferation marker Ki-67 is found extensively in HGSOC (Eckert et al., 2016), which inherently raises questions for a CSC model. If the proliferation of the tumour is driven from

only a handful of CSCs (or at most a larger pool of transit-amplifying cells) and the bulk of the tumour is quiescent, then why do entire HGSOC tumours express this proliferation marker?

Indeed, some studies have found that Ki-67 is actually higher in HGSOC than in low-grade serous ovarian tumours (Sallum et al., 2018), despite the latter being driven by upregulation of proliferative pathways rather than genomic instability. Potentially this may just represent aberrant upregulation of the gene as a result of genomic instability and dysregulation of the entire genome. Alternatively, it may indicate that proliferation is far more widespread than is predicted by a CSC model, and could be indicative of a tumour where the vast majority of cells are proliferative.

Studies have also investigated HGSOC tumours as populations, drawing crude approximations of the extent of proliferation from Ki-67 distribution (Chen et al., 2017). Chen *et al.* report an average of 40% of cells Ki-67$^+$ from 318 tumours, with some as high at 95%. This again appears inconsistent with a CSC model, and since each tumour is likely to show similar genomic instability, is not adequately explained by a hypothesis based on genomic copy-number driven aberrations. The data also show that low Ki-67 percentage correlates with chemoresistance and relapse. This is the opposite to what would be expected if the CSC hypothesis is true. If chemotherapy kills the bulk of cells leaving only a small proliferative chemoresistant subpopulation, then the expectation would be that the percentage of Ki-67+ cells would rise in resistant populations as the (proliferative) CSCs become a greater proportion of the remaining tumour, simply by attrition of the non-proliferative bulk cells. These observations are however consistent with a stochastic hypothesis in which more rapidly cycling cells are depleted more heavily, as they enter S-phase more often making them more available for ablation and giving them less time to repair DNA damage. This biased attrition would lead to a gradual increase in slow-cycling cells and a reduction in the percentage of Ki-67$^+$ cells. This premise is not even contingent on subpopulations with distinct proliferative rates, since the same process would apply to a normally distributed but homogenous population which would gradually shift to proliferate more slowly.

Apoptosis

Since it was first described in 1972 (Kerr et al., 1972), many studies have looked at apoptosis. These studies first elucidated the key proteins which act as the 'switching-mechanism' for the process, a large proportion of the family of cysteine-aspartate-specific proteases known as 'caspases', before cataloguing the plethora of pro- and anti-apoptotic factors whose interaction and competition regulates the process (Thornberry, 1998). In addition to this intracellular mechanistic detail, the web of associated signalling pathways and cell-cell interactions which potentially influence the process have been elucidated and clarified (Ashkenazi, 1998).

However, where studies have investigated apoptosis in a quantitative fashion, they have

predominantly sought to demonstrate upregulation or downregulation of the process, in order to show that an intervention via a particular protein or pathway has a significant effect on the process; this is the case both for studies investigating the basic biology of the process, and those which seek to manipulate it as part of translational therapeutic research. Quantitative assessment of the process can be measured in many different ways, for example assessing cleaved apoptosis cascade proteins by western blot or immunohistochemistry staining, or through similar methods such as a TUNEL assay (Zhu et al., 2013). Alternatively, apoptosis may be scored by manual counting of apoptotic bodies based on simple histology. By using these methods on experimental and control samples in parallel, it is easy to demonstrate that a given stimulus results in an X-fold increase/decrease in either the level of a particular protein, or the proportion of cells undergoing a particular stage of the apoptosis cascade; this latter measurement is often referred to as an apoptotic index (Potten, 1996).

A significant increase in apoptosis is an excellent way of demonstrating the impact of a perturbed pathway (or the effects of a bioactive molecule) on cell death. However, it is less useful in the context of understanding, and crucially *quantifying*, routine apoptosis. Later in this work, quantitative models will be created from experimental data in order to better understand proliferation. In order to optimise these models, it is important to take account of the process of apoptosis, and understand how growth in cell numbers is counteracted by the rate of cell death. However, apoptotic index relates to the absolute *apoptosis rate* only as a function of the length of the measured process (Potten, 1996). If 2% of cells are undergoing apoptosis as measured by a TUNEL assay, and the time-window over which a cell marks positively in the assay is 24 h, then the absolute rate of apoptosis is 2% of cells per day. If this window is only 12 h however, then only half a day's worth of events are being seen, and the absolute rate of apoptosis per day would be 4% / day.

Unfortunately, there is very limited data on the measurement of absolute apoptosis rate in any tissue, quite aside from HGSOC specifically. Data presented in a 1996 review of apoptosis by C.S. Potten (Potten, 1996) based on cells in the intestinal crypt, makes a good attempt to estimate this value. This work measured apoptotic index by manual histological scoring of apoptotic bodies, and then aimed to quantify the detection window by measuring how quickly apoptotic bodies are washed out after a pulse of a toxic stimulus. After a pulse of radiation or hydroxyurea (a simple cytotoxic chemotherapeutic), the number of remaining apoptotic bodies was assessed periodically and the decay rate calculated by regression. This gave 'half-lives' for the wash-out of 12 h and 3.5 h respectively.

This aspect of the data is interesting as it indicates that even the source of apoptosis appears to have a significant impact on the length of the process, as measured by the detection time-window over which apoptotic bodies are visible (Potten, 1996). The paper suggests that the differential arises as a result of cell-cycle specificity, as while both agents disrupt DNA, hydroxyurea induces apoptosis in S-phase while radiation is cell-cycle independent. This is interesting not only as it suggests the length of the apoptotic process is highly variable, but because this apparent variability will be important in discussion later in this thesis.

It is worth noting that this paper adds the caveat that this method is limited technically by the inability to stop additional cells from entering apoptosis after the toxic pulse, and that this will artificially inflate the half-life for wash-out of apoptotic cells. This is a reasonable argument, and would result in a shallower decay rate for apoptotic bodies, leading to a higher half-life value. A small shift could arguably have quite a significant impact on the already shallow gradient reckoned to have a half-life of 12 h. However, a similar absolute background rate-of-replacement would have a minimal impact of the comparatively steep gradient in the 3.5 h example, suggesting that this figure is likely to be relatively accurate.

The author notes data from work in the developing rat kidney which instead blocked apoptosis and followed the wash-out rate of apoptotic bodies (and inherently lacks the technical issue previously discussed). This found rates of between 1 and 2 hours. However, this assumes both that the apoptosis inhibiting agent does not also impact the progression of the apoptosis process, and that developmental processes are equivalent to adult ones. It also assumes that processes in different tissue types are equivalent. This does not appear to be the case. For example, in the intestinal crypt, it appears that apoptosis is higher in specific stem-like populations (Potten, 1977) (Potten, 1992). The nature of these populations is elucidated in more detail in a more recent paper which suggests that this population represents a quiescent tissue stem cell population (Buczacki et al., 2013). Even the absolute rate of apoptosis appears to be highly dependent on a range of factors, with work showing that apoptosis rates (as measured by apoptotic indices) appear to be affected by such cryptic effects as circadian rhythm (Potten et al., 1977).

The academic literature presenting absolute quantitative assessments of the length of the apoptotic process is quite limited since such technical details are of limited academic value unless they are crucial to a larger conclusion. However, sample data on such technical minutiae can be commonly found within the supporting information published by manufacturers of relevant reagents. Since the length of the apoptotic process does not impact the efficacy of the product, there is limited reason to expect these data to be biased for commercial reasons. A review of commercial supporting literature surrounding protein-based reagents (i.e. measuring activation of the caspase proteins driving apoptosis) suggests a consensus for apoptotic window around 10 h. For example, a guide to appropriate timings when designing experiments to measure apoptosis published by manufacturer Promega, presents sample data measuring cumulative caspase-3/7 and caspase-8 activity (via pro-fluorescent/-luminescent substrates). These data show sustained growth in the signal induced by active caspase-3/7 executioner caspases over a period of at least 7 h (the point at which measurement ends), implying that this represents a minimum window for the apoptotic process. This experiment induces apoptosis in a population of cells, so it could be argued that this represents cells entering the process at different times resulting in a constant net growth of signal even through individual cells undergo complete apoptotic clearance much earlier. However, this argument is not compatible with the data also presented for caspase-8 in the same experiment, which is observed to peak after only 3 h before falling away; this implies that after 3-4 h the caspase-8 substrate is already being degraded faster than it can be generated by

cleavage, indicating that peak caspase activity has long since passed, further implying that the majority of the cell population initiated apoptosis early. These observations are consistent with the respective roles of caspase-8 (an initiator caspase which serves as a transient 'trigger' for apoptosis) and caspase-3/7 (an effector caspase which undertakes the bulk of the proteolysis required for the controlled deconstruction of the cell). The continued caspase-3/7 activity after 7 h also suggests that apoptosis is a protracted process. Another dataset generated by Essen Biosciences in support of their caspase-3/7 assays measures caspase activity under similar conditions under varying levels of apoptotic inducer (campothecin). As expected, low campothecin levels result in slow and continuous signal growth consistent will cells sporadically entering apoptosis. However higher concentrations result in steep convergent sigmoid curves, consistent with rapid and extensive apoptosis initiation. These curves plateau after 12-15 h, implying that caspase-3/7 is active for at least this period, and establishing this as a minimum for the length of the apoptotic process. It should be noted that both these experiments were conducted in tissue culture in cell-lines, and that these results are therefore constrained by the limits of a 2D *in vitro* model.

In conclusion, apoptosis rates (and crude indices) appear highly variable across tissue types, stage, apoptotic stimulus and a range of further varied factors. Experiments such as this are inevitably limited by the constraints of the model systems in which they are conducted. The ultimate purpose of this question for the work that will be presented in this thesis is to understand the impact of background-level apoptosis within HGSOC tumour cells. Given the apparent impact of the inducing stimulus on the length of apoptosis, it is unclear how well these experiments measuring apoptosis induction in swathes of cells would actually represent the process occurring when a single isolated cell undergoes apoptosis; cell-cell signalling, and bystander effects may have a host of impacts in these experiments that would not be seen when isolated cells enter apoptosis. However, without a practical way to induce and measure the process in a single cell (without using a highly artificial model which would itself introduce huge confounding effects) it is impossible to be certain of how these processes take place *in vivo* in established human tumours. As such caution is required in interpretation of any measurements in this area, and when using apoptotic data it is important to ensure that the values being used are appropriate for the situation in question, and that the inevitable uncertainty involved has been properly considered.

1.4 Previous research

Previous work in the Brenton Laboratory (the PhD project of Siru Virtanen) identified phenotypic subpopulations within HGSOC, and presented a putative cancer stem cell population distinguished by the epithelial surface adhesion marker EpCAM. Sorting of $EpCAM^+$ and $EpCAM^-$ cells from HGSOC patient samples followed by growth *in vitro* showed no proliferation from $EpCAM^+$ cells, while $EpCAM^-$ cells were able to proliferate, and give rise to spheroids or colonies in respective *in vitro* experiments (Figure 1.2, top panel). Fur-

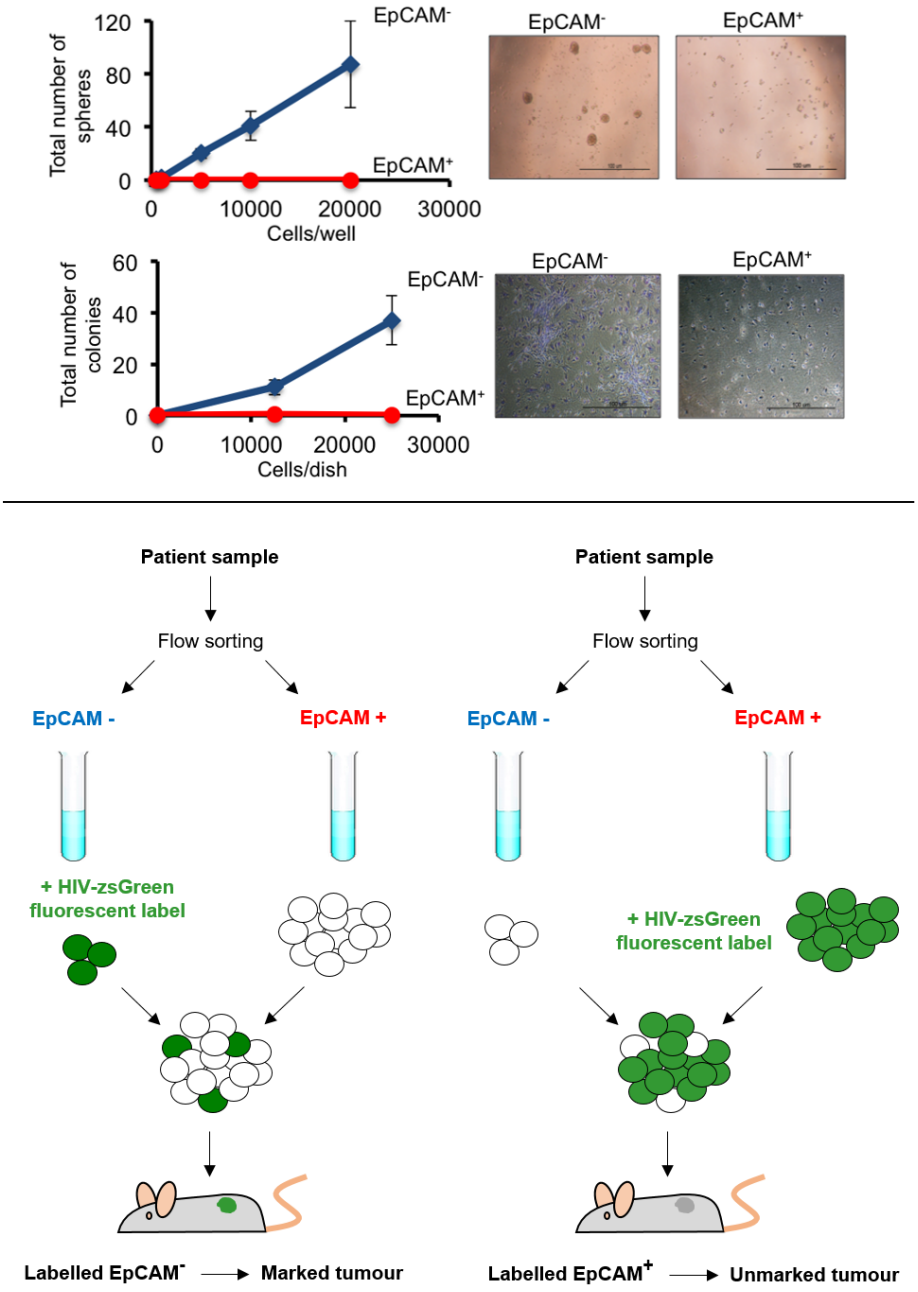
thermore, this EpCAM$^-$ subpopulation could be further subdivided by the marker surface marker podoplanin. The subpopulation positive for podoplanin showed the same ability to proliferate in *in vitro* assays, while the podoplanin negative subpopulation showed no proliferation, much like the EpCAM$^+$ subpopulation.

Attempts to replicate this experiment *in vivo* struggled at first, as none of the subpopulations appeared able to initiate tumours *in vivo* in isolation. In contrast, mixed cell populations were able to initiate tumours in the same context with far fewer cells. Based on the hypothesis that this result indicated that EpCAM$^-$ TICs are dependent on cell-cell interactions with bulk EpCAM$^+$ tumour cells, a second experiment was devised. Cells were sorted into subpopulations as before and, in each case, a single subpopulation labelled using a constitutive lentivirus expressing the fluorophore zsGreen (see Figure 1.2). When the EpCAM$^-$/Podoplanin$^+$ subpopulation was labelled, the majority of the resulting tumour showed green fluorescence labelling. Approximately 70% of the cells in the resultant tumour were labelled, which corresponds well with the ∼70% efficiency of the labelling process. When either the EpCAM$^-$/Podoplanin$^-$, or the EpCAM$^+$ subpopulation were labelled, the resulting tumour did not show detectable fluorescence. In all cases the resulting tumour was composed of a heterogeneous mix of all three subpopulations. This indicates that these *in vivo* tumours had arisen from the EpCAM$^-$/Podoplanin$^+$ subpopulation, and that this subpopulation had given rise to a phenotypically mixed tumour.

Given this evidence, a cancer stem cell (CSC) model was proposed (see Figure 1.3). This hypothesises that in a HGSOC tumour, the EpCAM$^-$/Podoplanin$^+$ cells represent a pool of indefinitely cycling CSCs which not only maintain and grow their own numbers, but are also the origin of the EpCAM$^+$ bulk tumour. The EpCAM$^+$ cells in this hypothesis are not proliferative, and their numbers are instead replenished by the EpCAM$^-$/Podoplanin$^+$ CSCs which can divide asymmetrically.

This replenishment may also occur through Transit-Amplifying Cells (TAs). In this variant on the CSC model, CSCs divide very slowly, giving rise to TA cells which can divide rapidly, but only for a limited number of cycles, before they themselves become quiescent. It is not clear what phenotype these TA cells would have since they would not be detected by the experiments previously described, since all proliferation is ultimately originating from the CSCs (since any transplanted TAs would be expended within a handful of division cycles).

One further observation made by Siru was that treatment of HGSOC cells with cisplatin chemotherapy both *in vitro* and *in vivo* appeared to selectively enrich for the TICs. This would suggest that these cells have some means of preferential survival under chemotherapy conditions. This finding might tend to favour a TA-based model, since this would allow for slow-cycling CSCs which would naturally be expected to be less affected by platinum chemotherapy. In this model the rapid division occurs within the TA fraction. These cells would be quickly killed off by chemotherapy, but could be repopulated by slow-growing CSCs. Alternatively, it is possible that the CSCs are faster cycling, but protected from chemotherapy by intrinsic resistance mechanisms (drug efflux or upregulated DNA repair



Adapted from "Phenotypic and Functional Characterisation of Cancer Stem Cells in Human High-Grade Serous Ovarian Carcinoma" (PhD thesis – Siru Virtanen)

Figure 1.2: Previous Brenton Laboratory experiments studying tumour initiation. Top panel shows *in vitro* sphere-forming and colony-forming assays. EpCAM$^-$ cells produce spheroids and colonies while EpCAM$^+$ do not appear to proliferate. Bottom panel shows labelling experiments in which each subpopulation in turn was labelled with an HIV-zsGreen constitutive marker. When the EpCAM$^-$ subpopulation was labelled, the resulting tumour was marked. When EpCAM$^+$ was labelled, the tumour was unmarked. Resulting spheroids, colonies, and xenografts all showed mixed heterogeneity, suggesting the EpCAM$^-$ subpopulation drives proliferation.

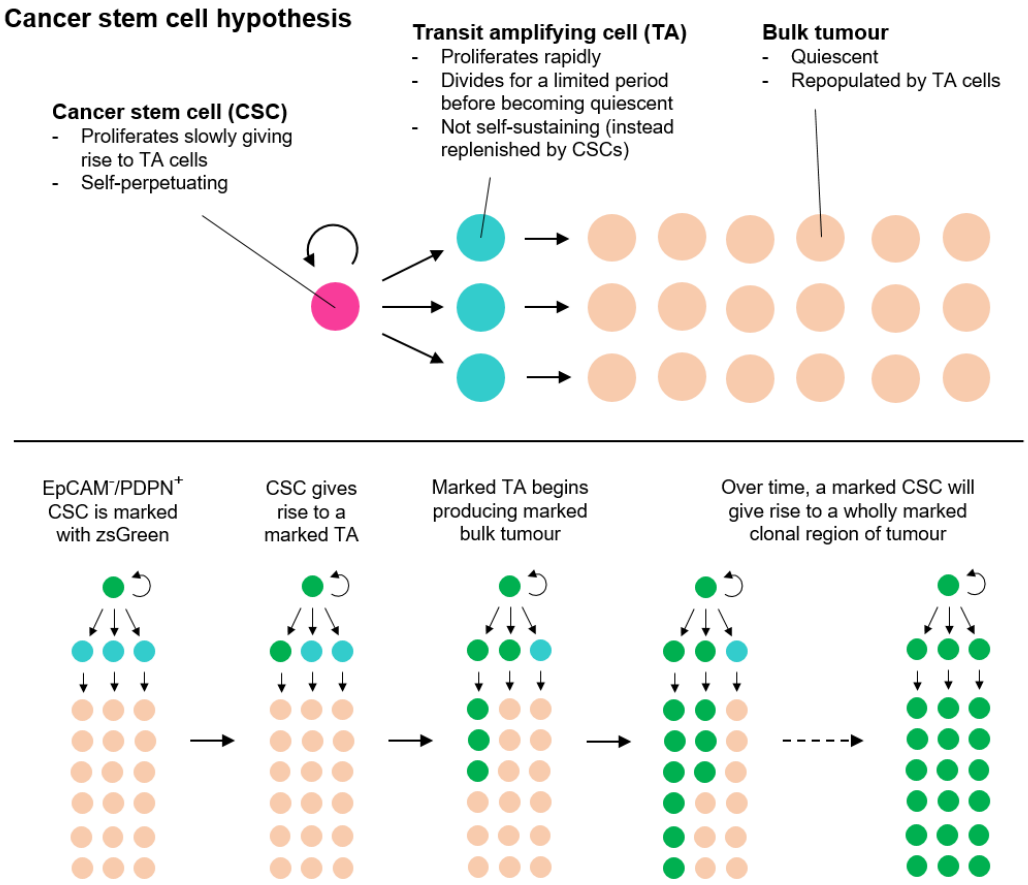


Figure 1.3: The cancer stem cell hypothesis. Top panel shows a model CSC hierarchy; bottom panel shows how such a hierarchy would explain the results summarised in Figure 1.2

mechanisms as discussed previously).

This was however a relatively minor experiment, and as a result the data are limited. In addition, although the EpCAM$^+$ fraction is much reduced, there is only a small increase in relative EpCAM$^-$/PDPN$^+$ proportion. This suggests that the bulk of the relative increase may be in the EpCAM$^-$/PDPN$^-$ fraction, which is not demonstrated to have tumour-initiating capacity.

In a 2015 paper, Janzen *et al.* reported a strong inverse correlation between CA-125 and TIC capacity (Janzen et al., 2015). The CA-125$^-$ subpopulation appears to be considerably more efficient in generating xenograft tumours (as well as resulting in larger tumour size and reduced latency). The conclusions of this paper are supported by previous literature suggesting that CA-125 may be downregulated in cells undergoing epithelial-to-mesenchymal transition (Comamala et al., 2011). Furthermore, this may explain why CA-125, used as a clinical marker for diagnosis of HGSOC, is suboptimal as a marker of relapse; if disease regrows from a chemoresistant subpopulation deficient in CA-125, this clinical test would only show a meaningful rise once these cells had given rise to bulk cells high in CA-125. Interestingly some tumours did still grow from the CA-125$^+$ fraction, though it is not clear if this represents CA-125 being imperfect for discriminating TICs, or simply the technical limits of separation within the experiment. The authors performed equivalent viral labelling experiments to those performed by Siru, and produced similar but not identical results. When the CA-125$^-$ subpopulation was labelled, the resulting tumours were heterogeneous but highly labelled, and similar conclusions can be drawn from this as can be drawn from Siru's work, suggesting that the CA-125$^-$ subpopulation and the EpCAM$^-$ subpopulation identified by Siru may be one and the same. However, when they labelled the CA-125$^+$ subpopulation, the resulting tumours were far less strongly labelled, and predominantly in only the CA-125$^+$ subpopulation, suggesting they may have successfully marked unipotent progenitors. This CA-125$^-$ population also shows enrichment in response to platinum therapy (again in accordance with Siru's work), and the authors build on this, demonstrating evidence to suggest upregulation of DDR factors, quicker resolution of γH2AX foci, and suppression of pro-apoptotic factors in favour of anti-apoptotic ones, including the cIAP proteins which inhibit caspase activation (although the rigour of the transcriptional data has been called into question as errors in the analysis have some to light (Janzen et al., 2016)). Furthermore they then go on to test the cIAP inhibiting agent Birinapant which when combined with platinum therapy shows a potent ability to deplete HGSOC xenograft tumours. Treatment with platinum alone results in a residual CA-125$^-$ population, but co-treatment appears to eliminate this population too, eliminating near to all tumour mass after 3 weeks of treatment. No tumour regrowth was seen in these samples, versus singly treated xenografts which were all observed to relapse – however, data for this follow-up is shown for only 3-4 weeks, so this may well simply represent a slightly longer latency to relapse rather than complete elimination of the disease.

1.5 Unanswered questions

The results from Siru's project represent convincing evidence that tumour-initiation originates in the $EpCAM^-/PDPN^+$ tumour fraction. These cells' ability to generate spheroids *in vitro* and to regenerate a heterogeneous tumour in *in vivo* xenograft experiments, indicates an ability to establish tumours, and to give rise to lineages representing multiple phenotypic subpopulations. However, these tumour-initiation experiments ultimately represent only a snapshot of growth at the point of initiation, by indirectly identifying the cell of origin. What they do not do is give any clear picture of the mechanics of tumour growth once initiation has occurred.

Tumour-initiation occurs directly after xenotransplantation (or transplantation *in vitro*) and therefore in the presence of the most limited degree of cancer-associated cells and favourable microenvironmental niches that these tumour cells will experience. In addition, in an *in vivo* environment this is complicated by the innate immune responses which may occur when the skin has been punctured by a needle. As the tumour grows it can be expected to recruit supporting stroma and vasculature, and generate bulk tumour cells, however immediately post-implantation, the cells are in the least pathophysiological environment that they will experience. This raises questions as to just how representative the proliferation being measured in tumour-initiation experiments is.

It is possible that tumour initiation occurs only in the observed TIC subpopulation because they are CSCs and other phenotypes are quiescent. Alternatively, it may simply be that these cells are the only population hardy enough to survive the transplantation process. Indeed, limited evidence from work by Siru via immunohistochemistry for EpCAM and proliferation marker Ki-67 on patient material appears to suggest that in this in-situ context $EpCAM^+$ cells are actually fairly proliferative, raising questions as to how representative conclusions from tumour-initiating studies are in a clinical context.

If these are the only cells that are still alive (and under sufficiently low stress to have not become senescent) after transplantation, it would not be surprising that they are uniquely able to initiate and regrow tumours. Equally, the ability of these cells to regenerate multiple subpopulations could be an inherent property that occurs under normal circumstances, or could be an aberrant 'wound-healing' type response to the depletion of other phenotypic subcompartments, in a similar fashion to the aberrant bipotency reported by Van Keymeulen *et al.* in the mammary gland basal cells in response to the loss of the luminal lineage ((Van Keymeulen et al., 2011); discussed previously). Exceptional asymmetric division could occur because the lack of $EpCAM^+$ bulk tumour cells (communicated by some secreted messenger, or cell-cell interaction) initiates a wound-healing response allowing the TICs to give rise to cells outside their normal lineage, or exceptional division asymmetry might be induced by external factors (i.e. resulting from host cells) arising from innate immune responses or tissue remodelling responses triggered as a result of the implantation process and tissue disruption.

It is very clear from these experiments that the identified TICs *can* give rise to multiple lineages and generate a heterogeneous tumour from single cells; a more pertinent question might be *do* they do so within an established tumour? The value of such research is ultimately rooted in its utility in disease management. In order to understand the nature of the tumours within HGSOC patients, the interactions between different HGSOC subpopulations must be studied in *established* disease rather than in the setting of initiation/implantation. This is not to say that tumour-initiation experiments do not have significant value – they may recapitulate well the environment of isolated tumour cells in the presence of a minimum of niche support, which may occur upon depletion of the tumour bulk through surgery or platinum chemotherapy. However, they only provide half of the picture. This project was designed to fill in the missing half of the picture and examine proliferation and hierarchy in *established tumours*.

1.6 Aims of this project

The purpose of this project is to investigate phenotypic heterogeneity within HGSOC within the broad context of translational research. Determining the characteristics of phenotypic subpopulations, particularly those that form the MRD population, would inform future therapeutic approaches and potentially allow drug design, or screening. An example would be allowing the development of a drug that could target the MRD subpopulation – this could be used to eliminate the MRD subpopulation before standard carboplatin chemotherapy was used to clear the bulk disease. However, if it is the case that the resistant population can be quickly repopulated by the bulk tumour, this approach would not be effective. In this case it might be more effective to instead develop a drug which exploits this plasticity, providing an artificial signal to promote interconversion of resistant cells into bulk tumour in the days leading up to carboplatin treatment. This would deplete the resistant subpopulation, potentially to a sifficient degree that subsequent platinum chemotherapy would completely eliminate the disease. In all the above cases, development of such drugs is dependent on a better understanding of the disease at the level of phenotypic heterogeneity within the established tumour, rather than in the context of the recapitulation of the disease from a small number of TICs.

The key to understanding phenotypic heterogeneity under these conditions is characterisation of the interactions between the cell subpopulations, since the objective is ultimately the identification of weak links which can be targeted by intervention. Since TIC capacity within the $EpCAM^-$ population has already been extensively and reliably demonstrated, further experiments need to determine whether this regeneration of the tumour from the $EpCAM^-$ population is mirrored in intact tumour, or whether it is simply an artefact of the *in vitro* or *in vivo* transplantation process (and the death or senescence of the vast majority of cells). A CSC model (left panel, Figure 1.4) would predict that this is representative of the normal tumour, and would predict the $EpCAM^-$ population constantly regenerat-

ing the tumour under both the stress of transplant, and also under typical tumour growth. However, the working hypothesis of this project is that this only holds true during implantation/engraftment due to selective elimination/senescence of other populations due to stress. This Dual Compartment hypothesis (right panel, Figure 1.4) conjectures that growth of the bulk EpCAM$^+$ subpopulation during *in situ* tumour growth comes from the bulk EpCAM$^+$ subpopulation itself, and not from bipotent TICs.

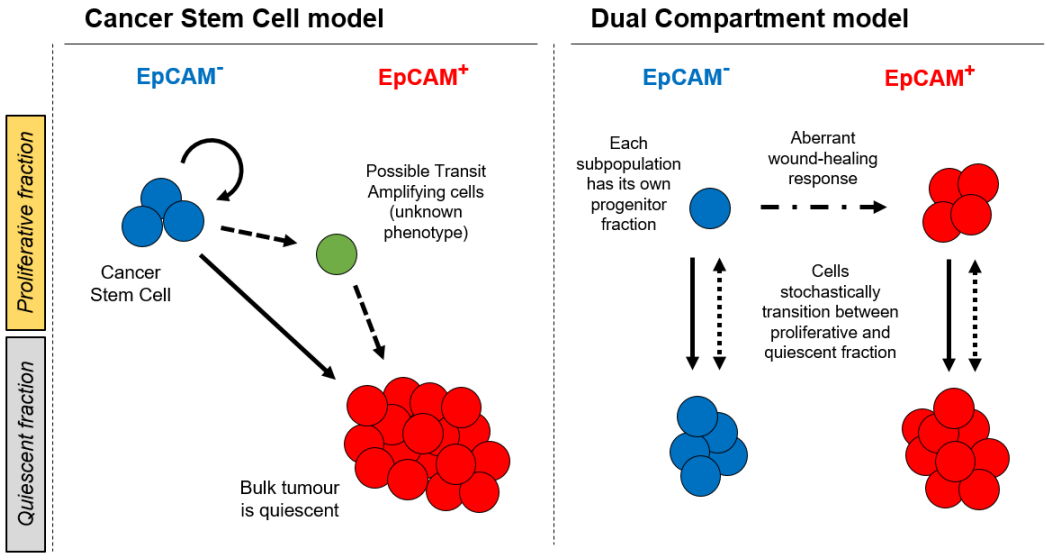


Figure 1.4: Visual abstract summarising the two archetypical hypotheses – the Cancer Stem Cell (Hierarchical) model, and the Dual Compartment (Stochastic) model.

To determine if this hypothesis is correct, two different methodological approaches are proposed. All require the tumour to be marked in some way, allowed to grow normally, and then that it be dissociated and assayed. The need to pre-treat tumours with tracing agents prevents direct study of patient material, since these cannot be introduced while the tumour is still in the patient, and patient tumours do not survive well *ex vivo*. This requires a model system to be developed that mimics the *in situ* tumour, but which can be experimentally manipulated. As such, patient derived xenografts (PDXs) were developed for experimentation (see chapter 3), through subcutaneous implantation of fresh HGSOC material from patients into NSG mice. These PDXs were then validated to confirm HGSOC status and demonstrate that they were representative of the original patient tumour. In addition, the phenotypic makeup of these populations was assessed, and compared to the results obtained in previous work (see chapter 4).

Two main experimental avenues were then pursued. The first was to interrogate rates of proliferation in each subpopulation in an *in situ* context (see chapter 5). This was via addition of nucleoside analogues, with proliferation in each phenotypic subpopulation

being determined by its rate of uptake of the marker. By performing this experiment over different timescales, it is possible to distinguish not only the rate at which subpopulations are expanding, but also whether the subpopulation itself is responsible for that proliferation, or if it is being repopulated by another subpopulation. The second was to investigate hierarchy within the tumour subpopulations in order to determine which subpopulations could give rise to which phenotypic lineages under the conditions of an established tumour (see chapter 6). The intention was to investigate the fate of HGSOC cells by using inducible lineage tracing constructs. By introducing these into a high percentage of tumour cells prior to implantation, then inducing at very low level once a tumour was established, it would be possible to mark individual cells at low density. By assessing the resulting clonal patches for their phenotypic makeup, hierarchy can be deduced based on the prevalence of mixed-lineage clones, and the degree to which single-lineage clones are observed. A CSC model would predict predominantly mixed clones, and no multicellular clones composed of $EpCAM^+$ cells only, while a model with distinct lineages would predict no clones containing both phenotypes (within the measurement error of the experiment).

1.7 Impact of this study

At present, the greatest challenge in the field of translational research into HGSOC is the limited numbers of tractable therapeutic approaches. The deployment of PARP inhibitor drugs in routine clinical practice in recent years represents a notable success story and an example of how an improved understanding of fundamental biology can be exploited to the clinical benefit of patients (Konecny & Kristeleit, 2016). However, these developments are only of benefit for a small percentage of patients, and further work is urgently required to improve the dismal prospects of the majority of HGSOC sufferers (Bowtell et al., 2015).

A number of promising concepts have been proposed which may offer additional angles for development of new therapeutic approaches. These include various approaches to stratifying HGSOC by genomic features, an approach which has already proven its value in breast cancer (Curtis et al., 2012). This is more complex in HGSOC as the disease is not driven by point mutations but rather by more complex genomic copy-number changes, which themselves are complex, diverse, and multifactorial (Mittempergher, 2016). Recent efforts are beginning to elucidate prognostic features within this complex 'genomic chaos', which with further development will provide additional tools to inform treatment. This type of work potentially opens up additional tractable therapeutic angles through the identification of dysfunctional pathways (e.g. PI3K–AKT signalling), defects in cellular processes (e.g. cell-cycle abnormalities) and routes to genomic dysfunction (e.g. breakage-fusion-bridge events) (Macintyre et al., 2018), which may be exploited for future translational advances.

Other efforts at stratification have focused not on categorising different tumours, but rather the cells found within each tumour population. The aim of these studies is to identify conserved forms of heterogeneity within these populations, and to determine how any sub-

populations interact in case any of these offer therapeutically tractable targets. The cancer stem cell (CSC) hypothesis is a particularly attractive one in this context as it proposes a clear hierarchy of cell types, with distinct proliferative characters, in which decapitation of the key cancer stem cell fraction would theoretically prevent disease regrowth, and would allow the remainder of the tumour to be eliminated with conventional chemotherapy. The attractiveness of the hypothesis is also increased by the convincing evidence of cancer stem cells in other malignant tissues such as glioblastoma (Chen et al., 2012), melanoma (Driessens et al., 2012), and intestinal adenoma (Kozar et al., 2013).

Therapeutic exploitation of cancer stem cells in HGSOC is dependent on developing a clear profile of these cells and then identifying exploitable targets for development of pharmacological intervention. As outlined extensively in subsection 1.3.3, research in this area has proved challenging, and a decade of studies have led to a plethora of proposed profiles, with many studies showing conflicting findings, and most finding that their proposed non-CSC fractions remain tumour initiating, albeit at a somewhat reduced rate.

This study sheds light on these findings by presenting data on proliferation dynamics within HGSOC tumours grown *in vivo*, along with empirical tumour growth rates, which indicate that proliferation within HGSOC is widespread across the tumour, and not isolated to a small fraction of cells. Examination of the parameters of these data through modelling of various hypothetical situations indicates that a CSC model would require significantly different parameters and suggests that the CSC hypothesis is wholly incompatible with the observations made here. This conclusion completely undermines the therapeutic rationale for interest in CSCs as it refutes the idea that there is a distinct subpopulation which if eliminated would render the tumour unable to regrow. As such it offers a basis on which to re-evaluate the focus of research in this area, rather than investing further time and resources in the isolation and characterisation of a subpopulation whose elusiveness now appears to be because it does not exist.

Beyond the generation of evidence to refute a prevailing hypothesis, this work also offers some initial clues into potential future avenues for study. The apparent stochastic nature of HGSOC proliferation outlined in this work, along with the slow-cylcing subpopulation identified, provides an elegantly simple explanation for disease reccurence observed in clinical practice, as the proportion of slow-cycling cells waxes and wanes in response to treatment and disease regrowth. This work indicates (although does not conclusively demonstrate) that the emergence of the slow-cycling population as the major population as a result of chemotherapy may well be due to the lower frequency at which these cells enter cell-cycle (as these cells are 'vulnerable' far less frequently). This outlines a possible biological vulnerability of the wider tumour cell population. If the rate of entry into cell-cycle is correlated with the rate of attrition as a result of chemotherapy, then further research into this area may prove clinically valuable. For example this hypothesis would suggest that extending the window over which patients are treated with carboplatin chemotherapy would kill a higher proportion of tumour cells and result in more effective debulking. Whether better debulking would

improve outcomes, and how the additional patient toxicity could be mitigated are open questions, but this research provides the necessary foundation for subsequent translational research.

Chapter 2

Materials and methods

The following list of materials and methods details final methodologies for each experiment described. Initial methodological approaches, and the optimisation process are described in more detail in later chapters.

2.1 Xenograft generation

2.1.1 Tissue processing

Patient solid tumour samples were obtained from Addenbrooke's Hospital as part of the OV04 study. Samples were collected by staff of the Addenbrooke's Hospital Human Research Tissue Bank who used a variety of methods, including imprinting methods and sections cut from samples snap-frozen in liquid nitrogen, to assess them (as discussed further in Chapter 3). Histological assessment and cellularity estimates were then (typically) made by a Consultant Pathologist. Where these samples were confirmed to be HGSOC, or where a complete assessment was not available, they were made available to the Brenton Lab. Confirmed non-HGSOC samples, and those determined to be benign were not used. Ascites samples (liquid build-up in the peritoneum containing single and clustered tumour cells) were also collected as part of the wider OV04 study.

Collected tumour material was processed according to a standard operating protocol established for handling of the OV04 solid tumour material. Samples were transported to the laboratory on ice, and processed as soon as possible upon receipt. Several small pieces ($0.25\,\mathrm{cm}^3$) were cut from each sample. The first was snap-frozen at $-80\,^{\circ}\mathrm{C}$, and stored for DNA extraction for genomic characterisation. The second was fixed for 24 hours in 10% neutral buffered formalin (NBF) before being transferred to 70% ethanol for wax embedding and sectioning (for histology and immunohistochemistry) within the institute's Histopathology

Core Facility. Samples were routinely haematoxylin/eosin stained to confirm a morphology consistent with tumour, and stained for p53 by immunohistochemistry (IHC) as exceptionally strong staining for p53 (due to over-stabilisation and aggregation of the protein) or else its complete abolition is a distinctive characteristic of HGSOC. Selected samples, including all of the tumours used for the *in vivo* work, were also further characterised, particularly for EpCAM and WT-1 (another distinctive marker of high-grade tumour). Additional tumour pieces were surgically implanted into mice as described in section 2.2, or were placed in freezing media and stored at −80 °C in the same fashion as for single cell solution (described below).

The remaining sample material was then mechanically dissociated (manually using scalpels) until the pieces were small enough to be drawn up in a 25 mm stripette, and then dissociated overnight (16 h) at 37 °C and 80 RPM on a rotary shaker in order to release the cells from any extracellular structures. Dissociation was conducted in a baffled flask using a dissociation mix of 14 ml DMEM:F12 (50:50) media (Gibco), 5 ml of 7.5% bovine serum albumin fraction V (Invitrogen), insulin (Sigma) at 2.5 ng ml^{-1} total concentration, gentamycin (Gibco) at 50 ng ml^{-1}, collagenase A (Roche) at 1 mg ml^{-1} and 100 U ml^{-1} hyaluronidase (Sigma). The resulting material was centrifuged at 475 g and the supernatant discarded (where samples with a high fat content gave rise to a layer of fatty material, this was pre-emptively aspirated off to avoid it coating the tube and the pelleted sample). Cells were washed with DMEM:F12 and centrifuged again at 210 g to remove non-cellular debris. The pellet was resuspended in 1 ml of 0.25% trypsin (in citrate buffer) (STEMCELL Technologies) for 4 min to break apart clusters of cells, then washed with DMEM:F12 and centrifuged at 475 g. Next the cells were resuspended in 1 ml of 5 U ml^{-1} dispase (in HBSS) (STEMCELL Technologies), containing DNAse (Sigma) to a total concentration of 0.1 mg ml^{-1}, for another 4 min to further separate cells, and degrade DNA from dead cells to prevent it binding live cells together; cells were then washed and pelleted as previously. Cells were resuspended in a pre-cooled (4 °C) 3:1 mix of 0.8% ammonium chloride solution (STEMCELL Technologies) and DMEM:F12 to lyse red blood cells, then washed and pelleted as previously. The pellet was resuspended in phosphate buffered saline (PBS) and passed through a 100 µm filter to remove undigested material. The cell number was estimated using a haemocytometer and a light microscope, and cells were resuspended in the appropriate volume of freezing media (50% heat-inactivated fetal bovine serum (FBS) (Gibco), 44% DMEM:F12, and 6% dimethyl sulphoxide (DMSO)(Fisher Chemical) and frozen at −80 °C in vials of 5 million or 10 million cells in volumes of 1 ml in 1.8 ml cryovials; freezing was performed in 'Mr Frosty' containers which use a propan-2-ol filled insulation layer to ensure a smooth temperature transition to maximise cell viability.

2.1.2 Xenograft processing

The protocol used to dissociate primary ovarian tumour samples was adapted and optimised for the dissociation of xenograft tumours. Xenograft bearing mice were killed by a Schedule 1

method (typically cervical dislocation) and their tumours removed using scalpel and forceps. Half of all the previously described volumes (of the same concentration) were used for each of the three enzymatic steps. Tumours were again mechanically dissociated with scalpels (xenograft tumours lack the fibrosis found in primary tumours, so this process is many times faster) and incubated in collagenase A / hyaluronidase dissociation mix for only 2 h. Given the tendency of the digested cells to clump (and possibly limit the enzymes' access to the cells) the sample was removed from the incubator after one hour and vigorously agitated by hand to break up clumps of material. Where the resulting cells were for immediate experimental use rather than culturing, gentamicin and insulin were not added. The centrifugation step at 210 g and the ammonium chloride step were omitted as xenografts have no fat, and minimal fibrosis and vascularisation, making these removal steps unnecessary.

2.1.3 Mechanical digestion

Xenograft tumours were removed as before and minced between two scalpels. In this case they were minced more extensively, with the aim of producing something with the consistency of paste rather than pieces $\sim$1 mm^3. Once minced, 1.5 ml of DMEM:F12 was added and the paste minced again. The liquid was then drawn up by pipette and filtered through a 100 µm filter to remove chunks and clumps. The resulting single-cell-solution was then counted, pelleted at 475 g for 5 min and resuspended in an appropriate volume. These cells were then used for subcutaneous implantation in the normal fashion (as described below).

2.2 *In vivo* work

Mouse work was approved by and conducted within the guidelines of the CRUK Cambridge Institute Animal Welfare and Ethics Review Board and the Home Office of the Government of the U.K. Mice used in this work were either bred in-house by the CRUK Cambridge Institute Biological Resources Unit, or purchased from Charles River Laboratories. Mice were fed and cleaned by the CRUK Cambridge Institute Biological Resources Unit. Two methods were used for xenografting of tumours, subcutaneous injection and subcutaneous surgical implantation. For injection, cells were dissociated as described previously, and resuspended in an appropriate volume of injection mix (50% DMEM:F12, 25% PBS, 25% growth-factor reduced matrigel (Becton Dickinson)). This material was injected subcutaneously into one or both flanks of a female NOD.Cg-Prkdc^{scid} Il2rg^{tm1WjI}/SzJ (NSG) mouse (BRU/Charles River Laboratories); each injection used between 50 µl and 100 µl of material containing between 1 and 10 (but typically 3) million cells. For subcutaneous surgical implantation, small pieces (approximately 2 mm in each dimension) of solid tumour (either from patient samples, or established PDXs) were cut. Mice were anaesthetised, treated with analgesics (carprofen at 6 mg kg^{-1} and vetergesic at 0.075 mg kg^{-1}, shaved, scrubbed with iodine disinfectant, and a small incision in the skin made. Tumour pieces were dipped in growth-factor

reduced matrigel and inserted under the skin, and the wound resealed with surgical glue.

2.3 Flow cytometry

2.3.1 Extracellular

Cells were obtained by the xenograft processing dissociation method discussed above, or from cell-lines grown in cell culture. Aliquots (not exceeding 5 million cells; typically 1 million cells) were transferred to flow cytometry tubes (Corning), resuspended in 100 µl of PBS containing the relevant antibodies at the concentrations detailed in Table 2.1, and incubated for 30 min at 4 °C in the dark. Cells were then washed with PBS and centrifuged at 475 g for 5 min. Cells were resuspended in PBS containing 1 µg ml^{-1} 4′,6-diamidino-2-phenylindole (DAPI), and analysed in 5 ml polystyrene round-bottom tubes (Corning). Most flow cytometry analysis for this work was conducted on an LSR II (Becton Dickinson); sorting was conducted on a FACS Aria (Becton Dickinson), analysis and sorting of unfixed unpassaged primary patient material using an INFLUX (Becton Dickinson), and some analysis of red and green fluorescent proteins (transfection and viral work) using a MACSQuant VYB (Miltenyi Biotech). For the INFLUX, polypropylene tubes (Corning) were used. Sorting on the Aria and INFLUX instruments was conducted by the CRUK Cambridge Institute Flow Cytometry Core Facility.

2.3.2 Intracellular

EdU was introduced *in vivo* by either intraperitoneal injection at 100 mg kg^{-1} (200 µl-350 µl at 10 mg ml^{-1}), or via supplementation of drinking water to a total concentration of 1 mg ml^{-1}. For initial experiments with cell lines, growth media were supplemented as per drinking water.

Intracellular flow cytometry was primarily used to assess incorporation of 5-Ethynyl-2′-deoxyuridine (EdU). This was done using the Click-iT Plus EdU Alexa Fluor 488 Imaging Kit (Invitrogen) using a slightly modified version of the provided protocol. All additional intracellular flow cytometry experiments used this same protocol and sample fixation/permeablisation using kit reagents. All centrifugation steps described were performed at 475 g for 5 min at 4 °C, and all incubation steps performed in the dark.
Cells were aliquoted as previously, but using 0.5 million cells per tube rather than 5 million. Since these cells must be permeablised to assay the intracellular environment, DAPI cannot be used as a viability dye; instead they were first incubated for 10 min at 4 °C with LIVE/DEAD Fixable Violet Dead Cell Stain (Invitrogen) at a 1:500 concentration (2 µl of fixable stain in 1 ml of PBS). Cells were washed with 1% BSA (from powder; Sigma) in PBS, centrifuged, resuspended in 100 µl of the proprietary fixative (provided with the kit)

and incubated for 15 min at room temperature. Cells were washed again with 1% BSA, centrifuged, resuspended in 100 µl of the proprietary permeablisation reagent provided in the kit diluted 10-fold in 1% BSA ('permeablisation wash') and incubated for 15 min at room temperature. The Click-iT cocktail was then made up as directed. The cell solution was aliquoted as appropriate and the Click-iT cocktail added to appropriate tubes. As excess buffer is not provided in the kit, PBS was added to control tubes. Cells were washed in permeablisation wash, centrifuged, and resuspended in 100 µl of BSA containing the appropriate combination of mixed antibodies (Table 2.1). These were incubated for 30 min at 4 °C. Cells were then washed in permeablisation wash, centrifuged, resuspended in 600 µl PBS supplemented with 1:150 propidium iodide and analysed/sorted by flow cytometry.

Flow Cytometry Antibodies				
Antibody	Manufacturer	Clone	Fluorophore	Final concentration
EpCAM	Miltenyi	HEA-125	APC	2 µl in 100 µl
CD90 (Thy1)	BioLegend	5E10	PE-Cy7	5 µl in 100 µl
Podoplanin	Biolegend	NC-08	PE	5 µl in 100 µl
CD10	BioLegend	Hl10a	APC-Cy7	5 µl in 100 µl
CD34	eBioscience	4H11	PE-Cy7	0.4 mg ml^{-1}
CD45	BioLegend	HI30	APC-Cy7	0.48 mg ml^{-1}
CD31	BioLegend	WM59	APC-Cy7	0.48 mg ml^{-1}
mBP-1	eBioscience	GC3	Biotin*	25 µg ml^{-1}
mCD34	BioLegend	HM34	Biotin*	25 µg ml^{-1}
mCD31	eBioscience	390	Biotin*	25 µg ml^{-1}
mCD45	eBioscience	30-F11	Biotin*	25 µg ml^{-1}
mTer-119	eBioscience	TER-119	Biotin*	25 µg ml^{-1}
mMHC C1 (H-2K^{b}/H-2D^{b})	BioLegend	28-8-6	Biotin	25 µg ml^{-1}
Biotin*	BioLegend	Streptavidin	APC-Cy7	10 µg ml^{-1}
HLA-ABC	Becton Dickinson	G46-2.6	BUV395	4 µl in 100 µl
CD298	BioLegend	LNH-94	PE	4 µl in 100 µl
p53	BioLegend	DO-7	FITC	1.24 µg ml^{-1}

Table 2.1: List of flow cytometry antibodies, and concentrations used. For antibodies where concentrations are not provided, dilutions are given. The prefix 'm' is used to indicate mouse antigens; all others are human. *Some antibodies were biotin conjugated and detected with the streptavidin secondary antibody listed.

2.4 Immunofluorescence

2.4.1 FFPE tissue sections

Where samples were pre-treated with BrdU, this was conducted in the same way as EdU (subsection 2.3.2). Tumours were surgically removed, fixed in NBF for 24 h, transferred to 70% ethanol, embedded in wax and sectioned at 3 µm thickness; embedding and sectioning

was performed by the CI Histology Core Facility. Sections were dewaxed using xylene, pure ethanol, and 70% ethanol washes using a Leica Technical Assistance Centre robot. Sections were boiled for 10 min in citrate buffer (10 mM citric acid, 0.05% Tween 20; pH 6) using a microwave, and allowed to cool. All following steps were performed at room temperature and incubations in the dark. Sections were washed with PBS, incubated for 5 min with 0.1% Triton X-100 (Fisher BioReagents), washed with PBS again and incubated for 1 h with 5% BSA (Sigma) strained through a 0.22 µm filter (Millex). Sections were then incubated for a further 1 h with appropriate primary antibodies, at the concentrations shown in Table 2.2, in 5% BSA solution. Sections were then incubated three times with PBS for 5 min to remove unbound primary antibodies. Secondary antibodies were then applied for a further 1 h as per Table 2.2. Sections were again incubated three times with PBS for 5 min to remove unbound antibodies, then incubated with DAPI at 1:1000 concentration (this solution being vigorously mixed with a vortex-mixer and filtered to remove possible precipitates). Sections were finally washed twice with PBS, allowed to dry for 10 min, then mounted using a coverslip and ProLong Diamond antifade mountant (Invitrogen). Sections were allowed to dry overnight and then imaged using an Axio Scan.Z1 (Zeiss). Where detection of multiple targets was performed on a single slide, this was done in parallel, with all primary antibodies being added at once, and the same with secondaries.

2.4.2 Single cell cytospins

Tumour cells, or sorted subpopulations, were resuspended at a concentration of 0.5 million cells per millilitre. Critically these cells must not be pre-fixed or they do not adhere, so flow sorting must be on extracellular targets only. These solutions were gently but thoroughly mixed by pipette, 200 µl (containing approximately 100,000 cells) added to White EZ Single Cytofunnels (Thermo Fisher Scientific) along with blank polysine slides (VWR), and centrifuged using a Cytospin 4 (Thermo Fisher Scientific) for 2 min at 450 g (2000 rpm) to give slides with small patches of cells adhered to the centre. These slides were fixed with pre-chilled (4 °C) methanol for 5 min at room temperature (methanol being added to the edges and allowed to flow across the centre, rather than being dropped directly onto the cells in order to avoid dislodging them), and the slides left in PBS at 4 °C overnight. Next day, immunofluorescence staining was performed as described above, beginning from the permeablization using Triton X-100. Imaging was performed using a Leica SP5 confocal microscope with HCX PL APO 63X/1.40 oil objective lens.

Primary Antibodies				
Target	Clone	Raised in	Manufacturer	Concentration
BrdU	B44	Mouse	Becton Dickinson	1:100
p53	DO-7	Mouse	Thermo Fisher	0.5 µg ml^{-1}
EpCAM	VU-1D9	Mouse	Abcam	1:100
EpCAM	Polyclonal	Rabbit	Atlas Antibodies	2 µg ml^{-1}
BrdU-AF555	3D4	Mouse	Becton Dickinson	1:5-1:10
Vimentin	EPR868(2)	Rabbit	Abcam	12.4 µg ml^{-1}
Vimentin-AF488	V9	Mouse	Abcam	1:100
BrdU	ICR1	Rat	Abcam	1:200
Secondary Antibodies				
Target	Fluorophore	Raised In	Manufacturer	Concentration
Mouse	AF647	Goat	Invitrogen	4 µg ml^{-1}
Rabbit	AF488	Goat	Invitrogen	4 µg ml^{-1}

Table 2.2: List of primary and secondary immunofluorescence antibodies, and concentrations used. All protein targets are human.

2.5 Transfection and viral methods

2.5.1 Transient transfection

Target cells were grown in tissue culture conditions. Cells were passaged into a new vessel 3 h before transfection, as partially-adhered cells appeared to take up plasmids more successfully. Cells were plated at 60% confluence in half the normal volume of media. The following transfection cocktail was prepared (in a 7 ml bijou) for transfection of a T225 flask (with a suitably scaled equivalent used for small vessels): 3 ml media, 27 µg plasmid DNA, 80 µl FuGENE HD transfection reagent (Promega). The cocktail was incubated for 15 min at room temperature, added dropwise to the cell culture vessel, and the vessel gently mixed and returned to the incubator. The transfection mix is not toxic to cells and so was only replaced as per routine maintenance of cell cultures.

2.5.2 Viral generation and transduction

Virus was generated using the transfection method described above in 293T cells. A suitable construct (15 µg; see subsection 6.3.1 for details of constructs used) was combined with the PAX.2 (7.5 µg and MD.2 (4.5 µg) second-generation viral packaging constructs (kindly provided by the laboratory of Greg Hannon) in media with FuGENE, and transfected as described previously. After 16 h the media (containing minimal virus) was discarded and replaced. After another 24 h, 48 h, and 72 h respectively, the viral-laden media was removed, replaced, and the viral solution stored at 4 °C. After the third collection the cells were disposed of.

The virus-laden media was pooled and ultracentrifuged in Thinwall Ultra-Clear 25 mm x 89 mm ultracentrifuge tubes (Beckman Coulter) using a SW 32 Ti rotor (Beckman Coulter)

at 80,000 g at 4 °C for 2 h. Viral solution was added to the tubes, underlaid with a 20% sucrose cushion, and topped up until almost full. After ultracentrifugation, the supernatant was poured off, and the inverted tubes allowed to stand for 3 min. The pellet was then resuspended overnight at 4 °C on a shaking incubator in 200 µl of 1% BSA, and then gently but extensively resuspended with a pipette then next day. Concentrated viral titre was either used immediately, or frozen at −80 °C.

2.6 Genomics and transcriptomics

Genomic and transcriptomic analysis was performed on snap-frozen tissue pieces stored from both patient samples and banked xenografts, as described previously. DNA extraction was performed on snap frozen tissue pieces using the DNeasy Blood & Tissue Kit (Qiagen). Patient tissues collected as part of my project (as well as the xenografts generated as part of my project), were also used in other work within the lab. I performed a range of DNA extractions for the work described, but many were also run by other Laboratory members in the course of their work.

2.6.1 STR genotyping

The genotyping was performed using the Promega PowerPlex 16HS Cell Line panel and resulting polmerase chain reaction (PCR) products separated using the Applied Biosystems ABI 3130 XL analyser using Pop 4 polymer (36 cm array). Sequencing for STR genotyping was then conducted by Genetica Laboratories. Fragments were then analysed (by the CRUK Cambridge Institute Biorepository) using Applied Biosystems Gene Mapper ID v3.2.1 software and the resulting profiles compared to published reference data. Percentage matches to test cells were then calculated. This process took advantage of the pre-established pipeline used within the CRUK Cambridge Institute for genotyping established cell-lines to verify identity.

2.6.2 Tagged Amplicon Sequencing (TAm-Seq)

Library preparation was performed by members of the Brenton Lab, and sequencing by the CRUK Cambridge Institute Genomics Core. TAm-Seq used the latest Optimal Panel of 697 primer pairs. This panel is an updated version of the Brenton Lab's Basic Panel used in previous work ((Forshew et al., 2012)), which now includes a number of additional genes and mechanisms of interest (discussed in subsection 3.4.3). An Access Array IFC (Fluidigm) microfluidics chip is loaded into its chip controller for priming, and then 50 ng genomic DNA samples are aliquoted in (in duplicates), along with the 48 pools of TAm-Seq primers. The Access Array IFC is then loaded back into its chip controller to run the microfluidics which

combine each possible primer set / DNA combination in a unique well of a 2304-well grid. The chip is then removed and run on a Fluidigm thermal cycler to PCR amplify the resulting gene products using the Fluidigm AA 48x48 Standard v1 protocol. The chip is then loaded back onto the controller in order to draw the PCR products back into the DNA loading wells. Dilutions of each sample-specific set of PCR products are combined with Access Array Barcode Library primers and put through 15 PCR cycles on a thermal cycler to generate barcoded libraries. The resulting libraries are assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies), normalised if needed, and pooled. AMPure XP Beads are used to clean up the library by removing small nucleoside fragments. Sequencing was then performed using paired-end 150 bp sequencing using the MiSeq system (Illumina). Sequencing data was analysed using a customized variant calling pipeline based on GATK haplotype caller. Key mutations or unexpected absences thereof were manually inspected using IGV (Integrative Genomics Viewer). This pipeline is summarised in previously published work (Piskorz et al., 2015).

2.6.3 Single cell mRNA sequencing

Single-cell mRNA sequencing was performed on eight PDX samples. As described previously, mice were killed, tumours removed, dissociated into single cell solution, and stained for EpCAM, CD90, and DAPI. These processes were all conducted at 4 °C (on ice, with pre-cooled equipment, and incubations in a fridge) and as rapidly as the methodology would allow. An unstained aliquot of cell solution was retained for the unsorted sample, and retained at 4 °C while sorting was completed. Samples were sorted using a BD INFLUX and 13,000 cells collected from each population. Tumour subpopulations were collected, resuspended in 100 µl of PBS, and a third of this solution used for library preparation. A similar number of unsorted cells were prepared at the same concentration.

Downstream library preparation and sequencing was then performed by the CRUK Cambridge Institute Genomics Core. Single-cell library preparation was performed on each subpopulation (and unsorted samples) using a Chromium Single Cell 3′ Kit (10x Genomics) according to the manufacturer's recommendations (Revision A), and cDNA concentration assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies) according to the manufacturer's recommendations. PCR steps were run for 12 and 12-14 cycles respectively, and size-selection on the Bioanalyser was between 200 bp and 9,000bp. Libraries were then pooled and sequenced.

Sequencing was performed using a HiSeq 4000 (Illumina) using paired-end Illumina sequencing technology. Asymmetric extension cycles (26,8,0,98) were used to cover read barcodes, sample index, and transcript insert respectively.

2.7 Miscellaneous Methods

2.7.1 Western Blotting

Cells for western blotting were grown in multiwell plates. Cells were dissociated using trypsin, centrifuged at 475 g at 4 °C and resuspended in Pierce RIPA buffer (Thermo Fisher Scientific). The samples were then sonicated for 3x 10 s in a water-bath sonicatior and returned to ice. Samples were mixed on a rotary shaker for 30 min and centrifuged at 21,000 g for 20 min to pellet cellular debris. The supernatant of protein lysate was aspirated and frozen. Frozen protein lysate was quantified using the Direct Detect system (Millipore), and the samples diluted as needed to normalise protein concentration in line with the lowest concentration (1-1.5 mg ml^{-1}).
For each sample, 32.5 µl of lysate was mixed with 12.5 µl NuPage LDS Sample Buffer (Invitrogen) and 5 µl NuPage Sample Reducing Buffer (Invitrogen) and incubated for 10 min at 70 °C in a heat block. Each lane of a NuPage 4-12% Bis-Tris Gel (Invitrogen) was loaded with 20 µl a given sample or with Precision Plus Protein All Blue Standard (Bio Rad) as a protein ladder. The gel was run for 30 min at 60 V (20 W) then 110 min at 120 V (20 W) in 1x NuPage MOPS Buffer (Invitrogen). The gel was then removed and blotted using an iBlot NC Mini Stack and the iBlot 2 dry-blotting instrument using preset Programme 0. The blotting membrane was then removed, trimmed, and washed in TBS buffer. Membrane blocking was performed with 1:1 Odyssey Blocking Buffer (LI-COR) and TBS for 1 h with gentle rocking at room temperature. Primary antibodies (see Table 2.3) were added at the concentrations shown in 1:1 Odyssey Blocking Buffer and TBS supplemented with 1:1000 Tween-20 overnight at 4 °C with gentle rocking. The membrane was then washed four times for 5 min with vigorous agitation in TBS supplemented with 1:1000 Tween-20. Secondary antibodies (Table 2.3) were added in 1:1 Odyssey Blocking Buffer and TBS supplemented with 1:1000 Tween-20 and incubated for 40 min at room temperature with gentle agitation. The membrane was then washed as previously, with the last wash using TBS only. The membrane was then imaged using a LI-COR Odyssey CLx.

2.7.2 Light-sheet microscopy

Patient derived HGSOC organoid samples were obtained from within the Brenton Lab, and cultured embedded in droplets of Type 2, PathClear Cultrex Reduced Growth Factor Basement Membrane Extract (Amsbio) in organoid growth media in 6-well plates at 5% CO_2 and 37 °C. This organoid growth media is made up 1:1:2 WNT conditioned media, R-spondin conditioned media, and Advanced DMEM/F-12 (Invitrogen), supplemented with 12 mM HEPES (Invitrogen), 1x B27 (Gibco), 1x N2 (Gibco), 10 ng ml^{-1} EGF (Peprotech), 100 ng ml^{-1} Noggin (Peprotech), 100 ng ml^{-1} FGF10 (Peprotech), 1 mM nicotinamide (Sigma Aldrich), 9 μM Y27632 (AbMole), 0.5 μM SB431542 (Cambridge Biosciences), and 2 mM glutamine (Sigma Aldrich).

Primary Antibodies				
Target	Clone	Raised in	Manufacturer	Concentration
mCherry	Polyclonal	Rabbit	Thermo Fisher	$0.36\,\mu g\,ml^{-1}$
tetR	9G9	Mouse	Clontech	$1\,\mu g\,ml^{-1}$
eGFP	F56-6A 1.2.3	Mouse	Cell Signalling Technology	1:1000
β-tubulin	Polyclonal	Rabbit	Cell Signalling Technology	1:1000
Secondary Antibodies				
Target	Emission	Raised in	Manufacturer	Concentration
Mouse (Secondary)	680 nm	Goat	Li-Cor	1:1000
Rabbit (Secondary)	800 nm	Goat	Li-Cor	1:1000

Table 2.3: List of antibodies used for western blotting, and concentrations used. For antibodies where concentration is not publicised, dilutions are given.

Three days prior to experiment, 20 pieces of 3 mm diameter fluorinated ethylene propylene capillary tubing (Adtech) were cut to type ∼5 cm in length, and sterilised in 70% ethanol overnight. Two days prior to experiment, 2% agarose was prepared and poured into a 6-well plate to a depth of ∼7 mm and allowed to cool. The tubing was removed from ethanol (with any residual droplets in the tube shaken out or removed with a pipette) and allowed to dry in a sterile hood. Media was removed from the organoid cultures, and organoid containing droplets dissociated with cold media and gentle pipetting with a P1000 for 5 min. The solution (of ∼30,000 cells) was pelleted at 400 g, and resuspended (using pipette tips pre-cooled in PBS kept on ice) in 1 ml of cold growth-factor-reduced matrigel (Becton Dickinson). Each capillary tube in turn was firmly jammed onto the end of a disposable tip of a P200, and the pipette used to draw around ∼10 mm of the cooled matrigel solution into the tube. The tube was then removed from the tip. Each tube was then pressed (matrigel containing end first) into the 2% agarose, resulting in an agar plug. These capillaries were then filled with the same organoid growth media as before, using a needle to avoid creating air bubbles which block diffusion, and dropped into flow cytometry tubes (4 per tube). A little growth media was added to the bottom so that it could diffuse into the tube through the agar. These were incubated as before for 2 days.

Two hours prior to the experiment, the media was replaced with the same solution supplemented with campothecin at 2 μM and IncuCyte Caspase-3/7 Green Apoptosis Assay Reagent (Essen Bioscience) at 1:1000 concentration (5 μM final concentration). Tubes were then inserted into the Lightsheet Z.1 (Zeiss) and the instrument calibrated. The chamber of the instrument was loaded with the same treated media as the tubes. The 2 h delay is a product of the time taken for a rigorous setup of the instrument, and the challenge of finding a good organoid / set of organoids for imaging; it is not a required step and is described here primarily so it can be considered when assessing the data presented. The number of tubes prepared is similarly because of the challenges of growing organoids under these conditions; eight tubes each for two samples were prepared, and five of these tested on the Lightsheet before a promising field of organoids suitable for imaging was located.

2.8 Statistical Methods

Analysis and data-processing for this work was conducted using various software packages as appropriate. Flow cytometry was performed using the FlowJo platform. Data analysis and statistical assessment was performed using R, supplemented with packages from the Tidyverse collection. Bar plots show data as points with averages (means) as wide orthoganal bars. All error-bars shown represent 95% confidence intervals, calculated as two standard errors from the mean. Regressions were performed in R using linear modelling functions, either directly via the lm() function or summary() functions which perform the same calculations. In all cases, shaded areas represent the region bounded by 95% confidence intervals for the regression which are calculated as two standard errors from the mean. Where multiple regressions are performed, regressions and ranges are colour-coded. The grey traces in Figs. 5.13, 5.15, and 5.17 indicate *de facto* 95% confidence intervals for the empirical data shown in black, calculated by taking gradients from the 95% confidence intervals around the regression shown in the left-hand panel of Fig. 5.13 and inputting these into the model respectively rather than the gradient of the regression itself which was used to model the empirical data. All p-values shown were calculated using one-tailed Student t-tests against a null hypothesis that the respective groups are equivalent, except where explicitly stated otherwise. For the logarithmic difference calculations, the null hypothesis is that there is no trend, since identical groups should show no variation on the y-axis against the x-axis. As an example, in Figure 5.9 the p-values are calculated via a one-tailed Student t-test against the hypothesis that log-difference does not correlate with time. Specifically, the calculation assesses the likelihood that the observed gradient has arisen by chance arrangement of experimental noise under the assumption that y is independent of x, and that the line should in the absence of any uncertainty be horizontal. The analysis of the data shown in Figure 5.21 was conducted using Pearsons' and Kendal correlation tests as outlined in the text; these were two-sided as the direction of any correlation was not established. The correlation measures shown (r and *tau*) measure the strength of a correlation between the two variables (and its direction), while the associated p-values assess the chances of that outcome arising by chance.

Chapter 3

Xenograft development

3.1 Introduction

3.1.1 Merits of clinical samples

In order to assess the complexities of growth and hierarchy within HGSOC tumours, an experimental model is required. Ideally this model should be as similar to the original tumour as possible, retaining similar phenotypic characteristics to the greatest extent possible. The Brenton Laboratory's strong history of collaboration with surgical teams in Addenbrookes' Hospital, means the Laboratory is ideally placed to access tumour samples direct from HGSOC patients upon surgical debulking and removal. One option would be to study this material directly. However, direct experimentation on such samples has several major drawbacks:

- Study can only be observational; patient samples can be analysed, but cannot be pre-treated – this prevents the use of techniques such as lineage tracing and drug treatment
- Patient material is limited and can only be used for a small number of experiments – thus the same material cannot reliably be used across multiple experiments, preventing rigorous reproducible research
- Patient samples often contain significant intratumoural heterogeneity, so several pieces of tumour from the same sample may not be similar, limiting reproducibility even within the small tumour pieces available
- Patient tumour cells come intermeshed with a wealth of other non-cancerous tissues (haematopoietic, stromal, etc.); these are hard to isolate, and contaminate experimental data, in particular making it hard to find non-epithelial tumour subpopulations,

should these be present

3.1.2 Merits of model systems

To avoid these problems, tumour material must be grown artificially in the lab, either through the use of stable cell-lines, the growth of patient tumour cells *in vitro* as either adherent colonies or spheroids, or growth *in vivo* as patient derived xenografts (PDX). However, each of these systems each have their own drawbacks. Cell-lines have limited complexity, and often their characteristics are quite far removed from the tissue of origin. This is because the cells suffer a great deal of stress both in the transition from growth in the patient to growth in a plastic flask, and also in adapting to growth in the absence of the non-tumour stromal cells; this results in massive selection of cells and the survival of only a very small subset, which are then expanded. Furthermore, growth in this artificial *in vitro* environment appears to result in the loss of most of the phenotypic complexity under investigation. As such, this model is not suitable. Colony assays and spheroid assays provide greater complexity, the former using irradiated stromal cells to provide a supporting (but non-proliferative) niche, the latter providing a three-dimensional structure and increased cell-cell interactions, and both using a supplemented culture medium with a cocktail of growth factors and cytokines to better simulate the original tumour site. Both methods produce a much more representative environment compared to cell-lines, but do not fully recapitulate the complexity of the original tumour. In addition, both result in slow growth, and both methods struggle to maintain viability in proliferation over the long term, limiting their utility in expanding stocks of tumour material.

Patient derived xenografts (PDXs) provide the most complex type of model. With PDXs, tumour cells are implanted within an immune-deficient mouse, which is used as a host to provide a biological environment that simulates the original tumour site, allowing a tumour to grow with a minimum of adaptation to its new environment (and a minimum of associated stress). PDXs introduce the additional complexities and costs of *in vivo* work, and are unable to simulate tumour-immunological interactions, but provide the best available mimic of the original tumour environment. The mouse also provides a naturally homeostatic environment that is ideal for maintaining long-term growth and rapid expansion of tumour.

PDXs have a number of limitations, and may not fully recapitulate the complexity within *in situ* HGSOC tumours. This may come about as a result of selection of cells during the implantation process (i.e. do the subset of surviving implanted cells which give rise to the PDX appropriately reflect the original tumour?) or selection at the biological level (i.e. do the samples which successfully give rise to PDXs fully represent the disease or are there biases in which tumours are able to grow?) and may mean the samples undergoing study do not fully reflect wider HGSOC disease. There is an argument here that since advanced metastatic cells are predisposed to survival and proliferation in adverse environments, these biases will cause these cells to be overrepresented, making PDXs good models of this form

of disease but less reflective of earlier stage localised HGSOC. However, it is worth noting that this advanced metastatic disease is overrepresented clinically anyway, since the disease is typically detected at late stage, and in addition it is this disease which is the driver of mortality. As such it is arguable that this late-stage disease is the most translationally relevant. These factors should be considered in the interpretation of data.

There are also some limits in terms of possible immunological interactions which may not be observed in immunodeficient mice; however current mouse models offer a relatively sophisticated recapitulation of a limited immune system, so this limitation is not nearly as broad a concern as it was for work in previous decades. Finally, there is some limitation on PDXs as a result of the murine nature of the tissue surrounding the tumour. This may potentially impact cell-cell signalling if contact and secreted signalling molecules are subtly different in the two species. PDXs may also introduce complications in terms of differential stroma microenvironment or vascularisation (either due to differences from the mouse system, differences in the availability of these structures at heterotopic sites, or due because the transplanted tissue is not as physically embedded into anatomical structures as host tissue).

There are numerous potential ways these features could theoretically lead to divergence between the PDX and the primary human model. However, despite extensive validation, this project has identified no reliable evidence of meaningful divergence in the fundamental biology of the PDX model versus *in situ* disease. It is important for this, and future work, to be extremely vigilant for potential confounding factors on this front, however it is important to recall that all scientific studies are inherently limited to some degree, and that confounding factors are ultimately an occupational hazard of the field. Despite the inherently imperfect nature of PDXs, they represent the current gold-standard of perturbable experimental systems, since they recapitulate greater complexity than other models, but provide a tractable and experimentally useful system, something that patient-derived material cannot match as *in vivo* it cannot be subject to experiment and *ex-vivo* it cannot be sustained reliably.

3.1.3 Xenograft approach

Since the core of this project is the assessment of phenotypic complexity, the PDX method is attractive because it allows expansion of tumour material, and avoids the problems associated with direct study of clinical samples, while retaining the subtleties of the phenotypic diversity being studied as fully as is possible outside of the original tumour site. There are a wealth of possibilities with regard to the generation of xenografts, in terms of the choice of mice, implantation site, and various forms of supplementation.

There are a variety of mice types available for xenograft studies, and in this case NSG (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ) mice were used. NSGs represent the gold-standard of immune-deficient mice, improving upon the older NOD-SCID strain, but are not as sophisticated as the humanised mice that are now available. NOD-SCID mice have been previously shown to engraft tumours less well (Quintana et al., 2008) (presumably due to their resid-

ual immune capabilities) making them an inefficient option in this case. Humanised mice supplement the immune-deficient genotype with a primitive transplanted immune system of human origin. These mice have the advantage of potentially better modelling tumour growth in that they can reflect some tumour-immune interactions absent from the NSG model. However, they come with a host of downsides including a drastic increase in cost and a risk of tumour rejection upon implantation during experimental setup (thus wasting mice). Critically, the transplanted immune progenitors are only stable for a few months, meaning experiments must be conducted within this limited window. Since early evidence from this project showed that establishment of some PDX models could take up to a year, this makes this model entirely unsuitable for establishment of the type of PDXs used in this study.

There are several possibilities for implantation sites for PDX models. One option is orthotopic implantation; by engrafting cells into the mouse's ovarian tract, the cells should receive a set of growth conditions that optimally represent the patient tumour nice. However, this option has downsides, primarily that the resulting tumour grows within the mouse anatomy making it hard to assess externally. As tumour volume must be continually assessed in order to meet ethics requirements, this necessitates *in vivo* imaging, which in turn dictates that tumours be pre-marked with luciferase. This adds a level of complexity that is extremely undesirable at an optimisation stage, where the boundaries of the conditions under which a PDX will grow are still being determined. In addition, the physiology of the mouse ovarian tract is somewhat different from the human, meaning any attempt at othotopic implantation can only ever be approximate. Finally, owing to the nature of the site, ovarian tumours are rarely well localised meaning that the true orthotopic site is often hard to determine.

Another option is to implant tumour material subcutaneously between the dermal and underlying subcutaneous membrane. Allowing a tumour to grow under the skin does mean the tumour is growing in an environment that does not optimally recapitulate the original tumour site, however it does provide several advantages. Chief among these is that the tumour volume can be easily assessed by feel or callipers without even needing to restrain the animal. In addition, moving the tumour away from the abdomen minimises the risk of tumours infiltrating organs directly, or outgrowing the cavity causing them to crush surrounding anatomy. This reduces waste of mice, loss of samples, and eliminates significantly possible harm to the mice used.

For the purposes of this project, the logistical complexities and ethical challenges of performing orthotopic transplantation outweighed the potential advantages of the approach, and all implantation work was performed using the subcutaneous method.

Finally implantation can be supplemented by a variety of additional techniques to improve engraftment rates. For example, the addition of osmotic minipumps to release hormones conducive to tumour growth is sometimes used to promote the growth of subcutaneously grown breast tumour PDXs. Unlike breast cancer, HGSOC is not known to be strongly influenced by hormones. The disease stratifies to only a small degree by progesterone-receptor

status, and not significantly for oestrogen-receptor status (Sieh et al., 2013) suggesting these hormones do not play a major role in disease maintenance and growth (as opposed to predisposition where increased hormone cycles are associated with increased risk). As such, the addition of hormone minipumps was not considered worth the logistical limitations it would impose. As minipumps typically last a week or two they must be constantly replaced requiring a significant outlay of time for surgery; this makes it hard to maintain large groups of mice without extensive manpower. While a constant hormone cycle may not be necessary, only female NSG mice were used for PDXs, on the grounds that the general hormone mix in the bloodstream might possibly have an effect, and because there was no major disadvantage to doing so. Finally all implantations were supplemented with growth-factor reduced matrigel. This has proved to be a valuable supplement to establishment in Siru's previous *in vitro* work and was thus carried across to the PDX methodology in the expectation that its ability to help tumour cells settle viably on a plate would also apply to *in vivo*.

3.2 Aims

The objective of this part of the project was to generate a range of subcutaneous xenograft models, each representing a tumour from a different patient.
These models also needed to be characterised to demonstrate that they represent good recapitulations of the original patient tumour.

3.3 Methods

3.3.1 Scheme of work

A summary of the overall approach to implantation of tissue from patient surgery to first-generation mouse is laid out in Figure 3.1; full details of tumour processing and implantation procedures are described in chapter 2.
Tissue for this project was available as part of the previously established OV04 programme. This project however presented some new challenges that had not been an issue in previous work. This necessitated some empirical testing to understand the issues, followed by iterative adaptation of existing procedures. Figure 3.2 summarises the evolution of the standard protocol both within tissue collection pathway and during processing within the lab, the details of which are discussed below.

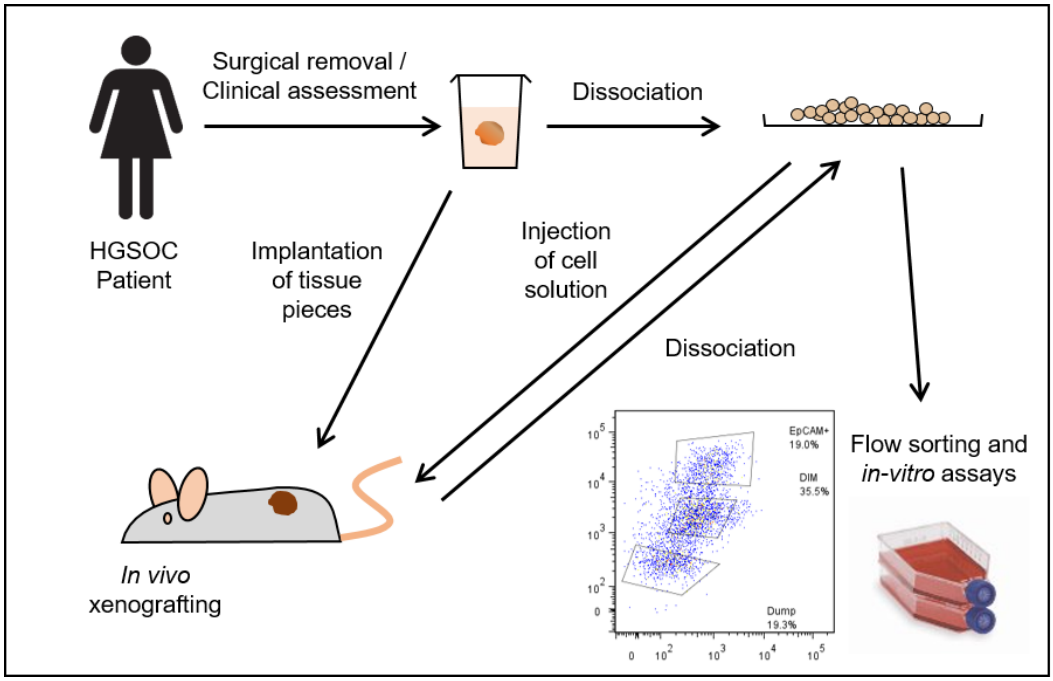


Figure 3.1: Summary of the overall process for generation of xenografts.

3.3.2 Iteration of implantation methodology for solid tumour

Initial method

During the first year of the project, tissue from suspected ovarian tumours was being collected by the Tissue Bank within Addenbrooke's hospital, material was taken as needed for clinical assessment, and any remaining material was made available to the Brenton Laboratory via the OV04 programme. The material would be transported to the lab, where (as per the protocol in chapter 2) it would be minced, digested overnight, and stored as a cell suspension frozen in liquid nitrogen. Frozen material would then be taken for assessment by flow cytometry and promising samples (i.e. those containing the largest proportion of apparent tumour cells) thawed, mixed with matrigel as described, and implanted subcutaneously. This approach was based on a small amount of similar work conducted previously in the Brenton Laboratory by Siru Virtanen, which had success with this methodology.
During this work, 26 OV04 samples were collected, prioritised based on *post-hoc* clinical assessment, and the best 5 chosen based on flow cytometry profiles. These 5 were then implanted by subcutaneous injection across 58 mice. No tumours grew. Due to logistical delays associated with Home Office approval of a personal licence for animal work, a quarter of these implantations were carried out by Siru; that these implantations also failed to grow tumours suggests inexperience or methodological error were not an issue. Additionally, subcutaneous implantations of cell-lines (PEO-1 and PEA-1) were also performed using an identical subcutaneous implantation protocol. While slow, these successfully grew tumours

in all cases, demonstrating no fundamental flaw in the methodology. Cell number was also considered, but despite increasing this from the initial 3M cells to 30M, there were still no successful tumour implantations.

Surgical implantation

One possible hypothesis to explain the failure of the previous method, is that the dissociation was too harsh, reducing the viability of the cells such that they are no longer sufficiently proliferative to be tumour initiating. In order to reduce this risk, samples were subsequently instead implanted surgically, rather than being injected as single cells. Pieces of tissue can be taken as soon as the tissue arrives at the lab, and implanted into mice within one to two hours. This avoids an overnight dissociation in collagenase/hyaluronidase, and processing using trypsin, all of which will reduce the viability of the cells. Immediate implantation also removes freezing and storage in liquid nitrogen, eliminating loss of cell viability at this step. Finally, by keeping tissues intact cell-cell interactions are maintained, ensuring that any niche-effects that may be present in the patient tumour are at least broadly maintained. Such a niche will degrade or change over time as human stromal cells are lost or replaced by murine stroma, but critically there will be minimal disruption during the period of peak stress: when the tumour cells are removed from their original environment and transplanted into their new host.

Tissue assessment by histological imprinting

The failure of a method previously shown to work, especially when that method fails in the hands of the scientist who developed and successfully used it, is perplexing. Having eliminated other major possibilities, the quality of the tissue arriving in the laboratory must be examined. Since clinical practice will adapt to new best-practice, logistical changes and pressures, and varying priorities and resource availability, it is quite possible that in the two to three years between these different sets of work, that the nature of the material arriving at the CRUK Cambridge Institute may be subtly different. A range of possible clinical factors could explain a reduction in the quality of available tissue: for example patients going for surgery may be being more heavily chemotherapy treated before receiving surgery, lowering the viable cell count; alternatively, surgeons may be being more aggressive in tumour resection, resulting in more non-tumour material being taken, reducing the tumour cellularity in the material reaching the lab. The switch from implantation using frozen cell solution to immediate implantation of material upon arrival, means that tissue cannot be assessed for quality before implantation into mice, as there will not be time to dissociate for flow assessment or wait for clinical assessment by the hospital. This is a challenge both in terms of cost and ethics, as it is far from ideal to be implanting every sample into mice when many will subsequently turn out to be tumours other than HGSOC (typically endometrioid or clear cell) or to be true HGSOC but at very low cellularity (5% or lower is not unusual).

As such, a method is required to allow samples to be assessed within the very limited timeframe of the surgical implantation protocol. After consideration and discussion with Addenbrooke's Tissue Bank, a methodology was agreed whereby at the stage of tissue being divided up and sent for clinical and research use respectively, histological imprinting was used to assess the crude histology of the sample. This method involves a tumour piece being pressed against a glass slide, leaving cells imprinted on the glass; these are then haematoxylin & eosin stained, and assessed by a pathologist (Kamatchi et al., 2015). While this does not allow an accurate assessment of the cellularity of the sample, or a watertight confirmation of the presence of HGSOC, it does allow a very rapid basic assessment of whether tumour is present in the sample, whether the histology of those cells is broadly consistent with HGSOC, and a rough approximation of the tumour density. This method provides a good foundation upon which to make a decision as to whether a sample is worth implanting.

Tissue assessment by frozen section

Based on the success of the imprinting method in providing advanced notice of which samples appear worth implanting *in vivo*, the protocol was further refined in order to provide more complete information up-front. By replacing the histological imprinting with assessment by frozen section (whereby a small piece of tumour is cut, snap-frozen at $-80\,^{\circ}\mathrm{C}$, and a section cut and stained with haematoxylin and eosin) several additional pieces of information are obtained; firstly, a more reliable idea of how consistent the sample is with HGSOC can be derived; secondly, drastically better approximation of the cellularity of the sample is gained; thirdly, a more accurate histological assessment can be made of a wider area of the sample, meaning that only the tumour-rich areas of the debulked material are dispatched for research use (and clinical assessment), saving time, money, and mice being wasted on minimal cellularity tissue. This method was adopted as a replacement for the imprinting process due to its multiple advantages (with imprinting as a reserve option as necessary, since it can be performed more quickly when manpower and time is limited).

3.3.3 Implantation of ascites samples

As part of the OV04 study, samples of ascitic fluid from the peritoneum, which contain large quantities of free-floating tumour cells, were also collected. Such fluid build up is common in HGSOC patients, and is drained as part of treatment. This material is of particular interest for research for several reasons. Firstly, it can be collected much less invasively than with surgery for solid material, meaning the procedure is common – this makes it both available, and also increases the chances of obtaining multiple sampling from the same patient, which is rare with surgery. Secondly, viable ascites cells have demonstrated the ability to survive in a free-floating environment. This indicates strong metastatic tendencies, and is indicative of advanced disease. These cells are therefore much more likely to survive the stresses of implantation and give rise to viable tumours. A number of these samples were implanted

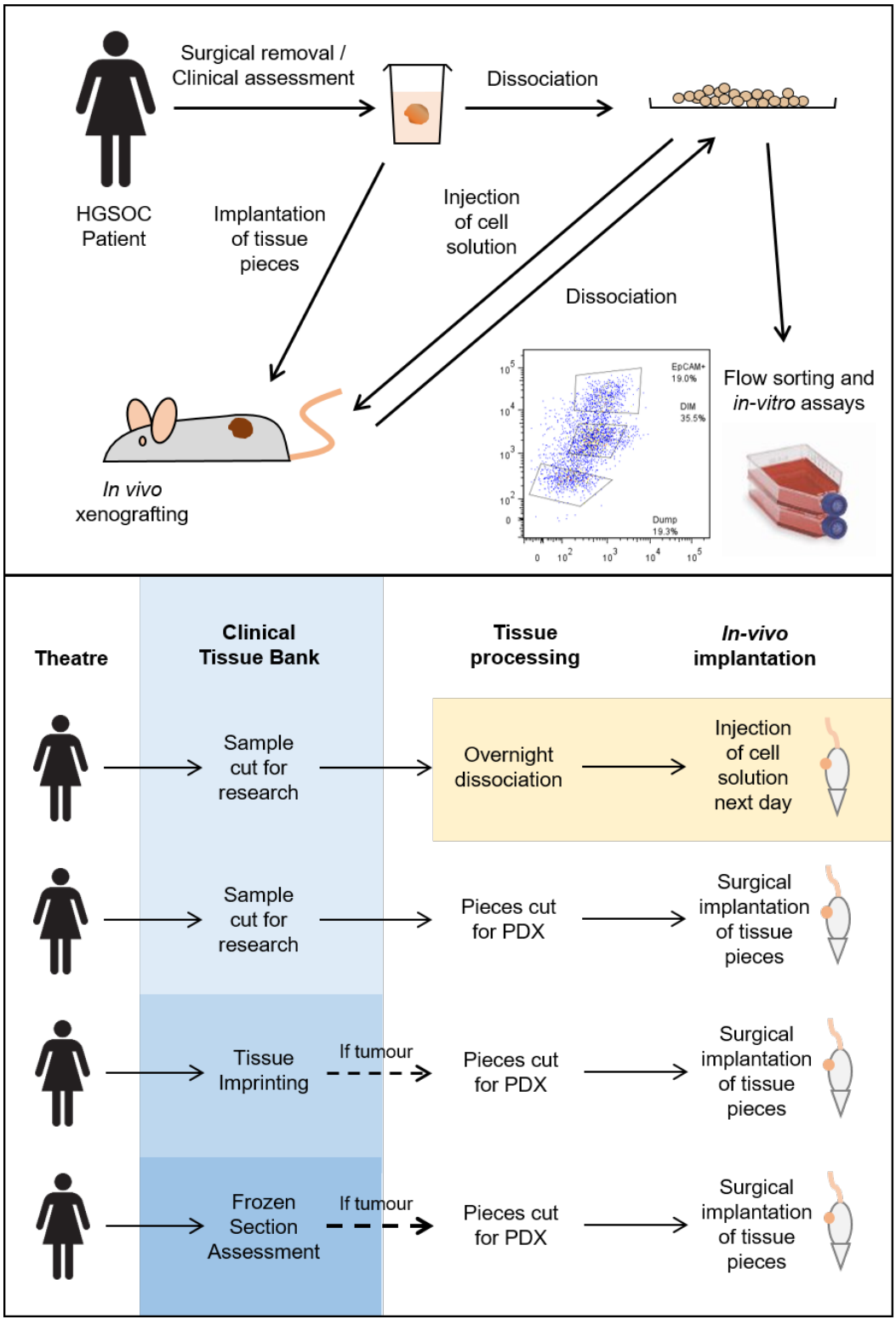


Figure 3.2: Summary of the evolution of method for assessment and selection of patient solid tumour samples, and subsequent implantation to form xenografts.

by injection (all from frozen stocks) to see if these might engraft, despite this method not being successful with the solid tumour samples. In addition, *in vitro* experiments by other members of the Laboratory have demonstrated differences between the single-cell fraction of the ascites, and the clumped spheroid fraction (and had optimised a filtration method). It has been suggested that spheroids may represent single stem-like cells which have proliferated in solution, creating clusters of daughter cells to form a spheroid. If this is the case, these may prove to be a better source of tumour initiation. To test this, the spheroid fraction was obtained from several of these samples, and injected in isolation.

3.4 Results

The outcomes of first-generation xenografting are summarised in Table 3.1.

First generation implantation success rates (by method)		
Method	Success by biological replicates	Success by technical replicates
Solid tumour (unassessed) by injection	0% (0/5)	0% (0/116)
Solid tumour (imprinted) by surgery	40% (2/5)	31% (5/16)
Solid tumour (frozen section) by surgery (poor cellularity)	9% (1/11)	5% (2/42)
Solid tumour (frozen section) by surgery (high cellularity)	100% (5/5)	45% (9/20)
Whole ascites	91% (10/11)	69% (29/42)
Ascites spheroids	43% (3/7)	25% (4/16)

Table 3.1: Success rates of implantation stratified by method used, and pre-implantation assessment where appropriate. Biological replicates refers to different distinct patient samples. Technical replicates refers to distinct implantations (mice typically received an injection in each flank, such that each mouse is treated as two replicates). High cellularity was defined as $\geq 30\%$.

As expected, improved success rates are achieved with each iteration of the method. While the relatively small sample means this does not represent a statistically rigorous proof of continuous improvement (as that was not the objective here), it is certainly indicative of improved implantation success, and the optimised method would represent a sound starting point for deriving PDXs if wider array of samples were needed, something that could not be said of the initial approach.

Ascites samples demonstrated effective engraftment, even when injected as single-cell solution. Not only this, they were able to demonstrate very high efficiency in implantation despite this disadvantage. It should be noted that a large number of stored ascites samples were available for implantation, and a significant amount of *in vitro* study had been performed as part of other work in the Brenton Lab. As such, a large pool of samples

were available for *in vivo* implantation, and only the best were selected, based on positive indications from previous studies. Therefore, it may be more appropriate to compare the success of these samples against the selected solid-tumour samples, rather than the initial unassessed ones. However, even in comparison to only the high cellularity solid samples, the ascites samples arguably outperform them, with the rate of biological replicates successfully engrafting at least once being almost as high (91% versus 100%), but with a noticeably higher rate of success at the level of technical replicates (69% versus 45%).

These values are broadly consistent with other reports from the published literature (detailed in subsection 1.3.4). Accurate assessment is challenging since many of these publications report success rates in terms of a single top-line figure and do not discriminate between technical and biological success rates. The Mayo Clinic's large PDX development programme reported creating 168 different biological models at 70% efficiency **??**. The work reported here compares well, as even the technical rate for injected samples is 69%, despite a significantly smaller programme and less opportunity for optimisation. However, efficiency rates in this work are well below the 90-100% efficiency (implied but not definitely the technical efficiency) consistently reported by the Neel Laboratory in their work **?? ??**. The authors of these papers use a variety of implantation sites at high efficiency, as well as both primary and relapse tumours, so their success is unlikely to be attributable to either of these key factors. Equally their methodology is very similar to the work presented here, and there are no differences which would explain the discrepancy. There are two immediately apparent potential explanations. The first is that this reported figure is actually for biological replicate efficiency, in which case it mirrors the 90-100% efficiency seen in this work. Alternatively, it is possible that differing approaches in the details of diagnosis and treatment of HGSOC between the UK and Canada result in tumour material with different properties which has a knock-on effect on the engraftment efficiency. The best comparator for the work outlined here would be the work done by Kryczek *et al.* who used similar subcutaneous implantation techniques, however they do not report implantation efficiencies, so this is unfortunately not possible (Kryczek et al., 2011).

Unexpectedly, purified ascites spheroids proved significantly less successful than the unpurified mixed samples composed of spheroids and single cells at both the level of biological and technical replicates. It was presumed that cells capable of proliferating in isolation (to give rise to a spheroid) would be more readily able to engraft to form a xenograft. One possible explanation would be that the presence of the spheroid actually supports the cells inside with cell-cell interactions and adhesion, and as such the proliferating cells are actually under weaker selection pressure than the isolated single ascites cells which are demonstrably able to survive alone. It is assumed that these spheroids arise from isolated cells, but perhaps these spheroids actually arise from aggregation, or perhaps the cells that gave rise to the spheroid, once surrounded by a new niche, and lose whatever 'edge' they had, meaning they (unlike the isolated cells) are no longer able to engraft once transplanted *in vivo*.

3.4.1 Serial passage

Subsequent experiments in this project require not just PDX models, but cohorts of mice implanted with similar tumours, which can be further subdivided and subject to controlled experimental conditions. In order for this approach to be valid, all mice must carry the same tumour, and these tumours must grow in approximately the same fashion. Practical reality also dictates that all tumours must grow at approximately the same rate, so that each arm of the experiment is being treated in parallel. Otherwise each mouse becomes a separate experiment and maintaining consistent conditions becomes near impossible (while meaningful trends start to become submerged in batch effects).

Initial design of the experimental work had assumed taking a fresh tumour, creating a homogenised cell solution, implanting it into 10-20 mice, allowing the tumours to grow, and then performing experiments on these batches. However, it soon became clear that, in light of initial experiments, this was not viable for three reasons:

- Even for tumour samples which implanted well, the technical success rate for first passage implantation was well below 100%; thus even where groups of tumour bearing mice were generated, the cohorts were underpowered. This could only be rectified by adding a significant excess of technical replicates to account for failures (given an optimistic 50% average technical success rate, double the number of mice per cohort would be needed to get a fully powered cohort only half of the time).
- Growth times for these first generation samples were extremely divergent (even within individual samples) as shown in Figure 3.3. Statistical analysis or extrapolation is dificult due to the large number of samples with only a single success, but of the replicates with multiple successful replicates per sample, half of the samples reach their limit over a range of at least two months. In a bigger cohort, this window is only going to grow, and this length of time already represents a problem as the quickest growing tumours will hit maximum permitted size before the slowest rowing tumours have begun to grow.
- Moving to a surgical implantation method for solid tumour raises issues with regard to generating a homogenous cohort. While injection of single cell solution into a cohort of mice does not guarantee this, it is at least reasonable to expect a roughly even distribution of any cellular heterogeneity across the injections, as the dissociation will result in effective mixing of any hetergeneity. With surgical implantation there is significant risk that cells are not evenly distributed across the original tumour, and it cannot be assumed that one piece of implanted patient material is representative of a second. As such there is potential for a high degree of variation to be introduced to the cohort, drastically weakening the experiment. It is not even necessary for phenotypic heterogeneity to be unevenly distributed; uneven distribution of tumour cellularity alone would more than account for any variation, as a piece of tumour that is predominantly tumour will grow faster than one composed mostly of stroma and

fibrosis, and cellularity cannot be accurately judged by eye.

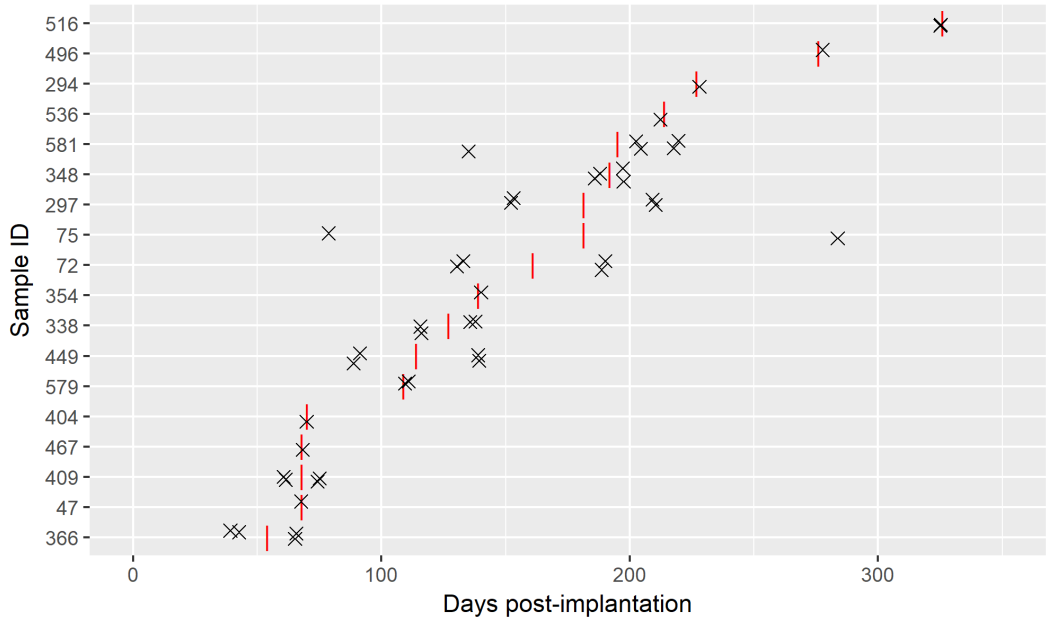


Figure 3.3: Summary of the time between implantation and tumour harvest for all first generation xenograft samples. Individual tumours are shown as crosses, and average lag-time by red vertical dashes.

In order to expand available stocks of PDX material for *in vitro* work, tumours that successfully grew in mice were serially-passaged mouse-to-mouse by the same method used for first-generation implantation. Interestingly, the resulting second-generation tumours were far less unpredictable, growing both much quicker, and more consistently. Passaged tumours (both at second and subsequent generation) grow in around 2-3 months, making their growth on par with that of the fastest growing samples. In some cohorts this period can be as long as 4 months for tumours derived from certain biological samples, but even here there is minimal variation in growth of tumours within the same cohort. Furthermore the success rate for passaged tumours is extremely good, and for samples that grew well in initial implantation (i.e. more than one successful tumour) it approaches 100%. Reliably knowing that most if not all samples in a cohort will grow is another critical requirement for practical experimental studies.

The samples that were successfully serial passaged are shown in Table 3.2. The broad improvement in apparent quality is likely to be due to improvements in engraftment allowing for more stringent selection of good samples, rather than a fundamental improvement in method or tissue quality. It is also likely to reflect the increased reliability observed with the implantation of ascites. A number of good quality samples that grew late on in the project were not passaged further, and so do not appear here, because there was already an

adequate bank of samples available, and generation of a larger stock was not a priority for this work.

List of passaged xenografts							
OV04	Type	*TP53* mut.	p53 protein	Age	Stage	Location	Notes
404	ST	None	W/T	60	3B	Liver	Endo. (histology)
536	ST	None	W/T	52	3A	Left ovary	Endo. (TAm-Seq)
496	ST	c.1024C>T	p.R342*	59	1C1	Abdomen	Unreliable passage
516	ST	c.742C>T	p.R248W	76	3C	Omentum	Unreliable passage
354	ST	c.759celC	p.I254Sfs*91	52	1C1	Sigmoid colon	Unreliable passage
449	ST	c.403T>G	p.C135G	59	2B	Colon	
579	ST	c.536A>G	p.H179R	66	4	Omentum	Unreliable passage
581	ST	c.743G>A	p.R248Q	78	3C	Omentum/uterus	
075	A	c.991C>T	p.Q331*	55	3C	peritoneum	
366	A	c.783delT	p.S261Rfs*84	51	4	peritoneum	
409	A	Not tested	Not tested	75	4	peritoneum	lethal *in vivo*
338	A	c.637C>T	p.R213*	38	4	peritoneum	
297	A	c.661G>T	p.E221*	72	4	peritoneum	
348	A	c.637C>T	p.R213*	58	4	peritoneum	
072	A	c.574C>T	p.R213*	59	4	peritoneum	

Table 3.2: List of xenografts considered for experimental use. Samples in red were rejected for the listed reasons. Samples in yellow were to some extent unreliable in their passaged growth, at least initially, and were only used once cohorts could be grown predictably. W/T = wild-type p53 protein; Endo. = endometrioid disease rather than HGSOC (text in brackets shows method by which this was determined). Types: ST = solid tumour; A = ascites. Age and stage refer to patient of origin.

The use of passaged samples does, however, raise a number of concerns surrounding how representative they are with regard to the first generation of PDX (and ultimately to the original patient sample). For the samples from which xenografts were successfully grown (shown in Table 3.2), validation was performed to confirm that they are not only HGSOC, but also representative of the original patient tumour.

3.4.2 Validation by histology

The first step in confirming HGSOC status is a more detailed histological assessment (since the initial clinical assessment is performed crudely from a single slide under time pressure). In addition to receiving information on the HGSOC status of the original tumour (from routine clinical assessment), the passaged PDX also needs to be assessed to confirm that the tumour that has arisen is actually HGSOC. Validation was performed by assessment of a combination of:

- H&E (haematoxylin/eosin to assess crude histology)

- p53 (either overexpressed or completely absent in HGSOC)
- WT-1 (a p53 interactor which is upregulated in HGSOC)
- EpCAM (to confirm epithelial status)

For some of the samples, particularly those from ascites, it was not possible to get a patient H&E, and for some of those shown this material may not represent the most viable part of the tumour sample. As seen in Figure 3.6, at least one of the passaged samples (OV04-404), assessed as HGSOC from frozen section, gave rise to a PDX line that is not consistent with HGSOC staining, demonstrating the importance of validating all PDX lines generated. The subsequent clinical assessment for this sample reported that it was mixed HGSOC / endometrioid, so the resulting xenograft is likely to be composed of the latter. Samples OV04-348 and -366 demonstrate concerningly equivocal p53 staining, but show strong WT1, suggesting that they are actually HGSOC. Sample OV04-354 shows very poor staining for p53, and does not appear to be HGSOC by histology. This sample would have been excluded, had subsequent validation methods not shown it to be consistent with HGSOC at the genome level. The uncertainty over the status of this sample (along with its erratic reliability in early passages) resulted in it not being used extensively in subsequent studies. Sample OV04-536 appears to be consistent with HGSOC given the complete absence of p53. However, subsequent validations demonstrate this sample is not HGSOC, but endometrioid. This underscores the limitations of histology, and need to validate such samples at deeper levels. Sample OV04-409 shows all indications of being HGSOC, however this sample was found to rapidly give rise to lethal metastases when implanted *in vivo* into mice. These metastases developed long before any subcutaneous tumours were observed, and the first indication of their presence was signs of severe sickness, or mortality, in the host mice. Upon necropsy, numerous tiny (though visible to the naked eye) tumour metastases were visible on the surface of the lungs and/or liver, and may have been present in other organs at microscopic levels. The lethal nature of this sample *in vivo* prohibits further *in vivo* work with it, due to both ethical considerations and that host mice would not survive long enough to reach experimentally useful endpoints. After the sudden death of one mouse, and euthanasia of another due to signs of severe sickness, the remaining hosts of this sample were euthanised and a moratorium put on its further *in vivo* use.

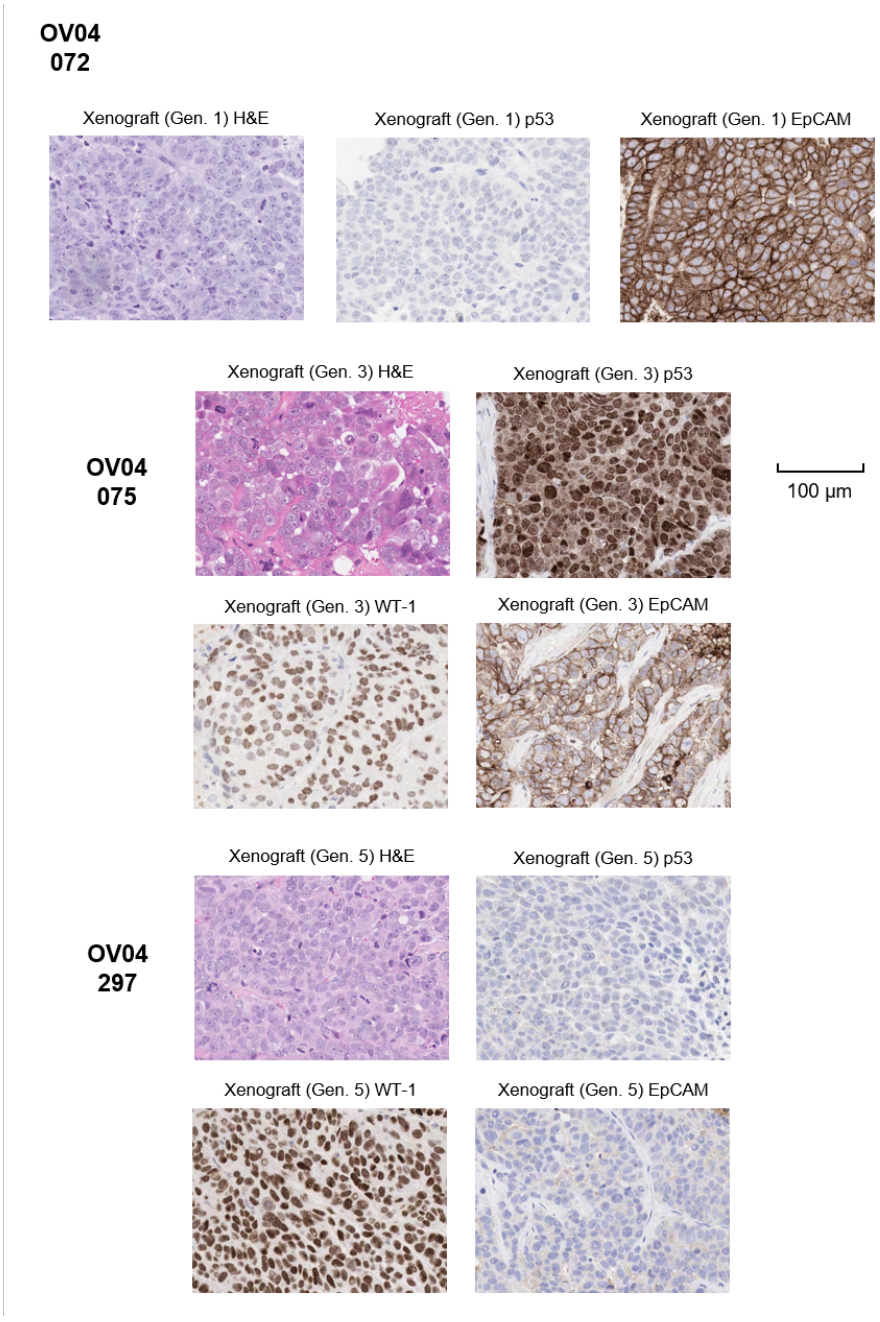


Figure 3.4: Validation of xenografts by histology (1/5). Haematoxylin and eosin (H&E) staining is shown for both patient sample (where available), and resulting PDXs. To confirm HGSOC status of the PDX, p53 and WT1 immunohistochemistry is also shown where available, along with EpCAM.

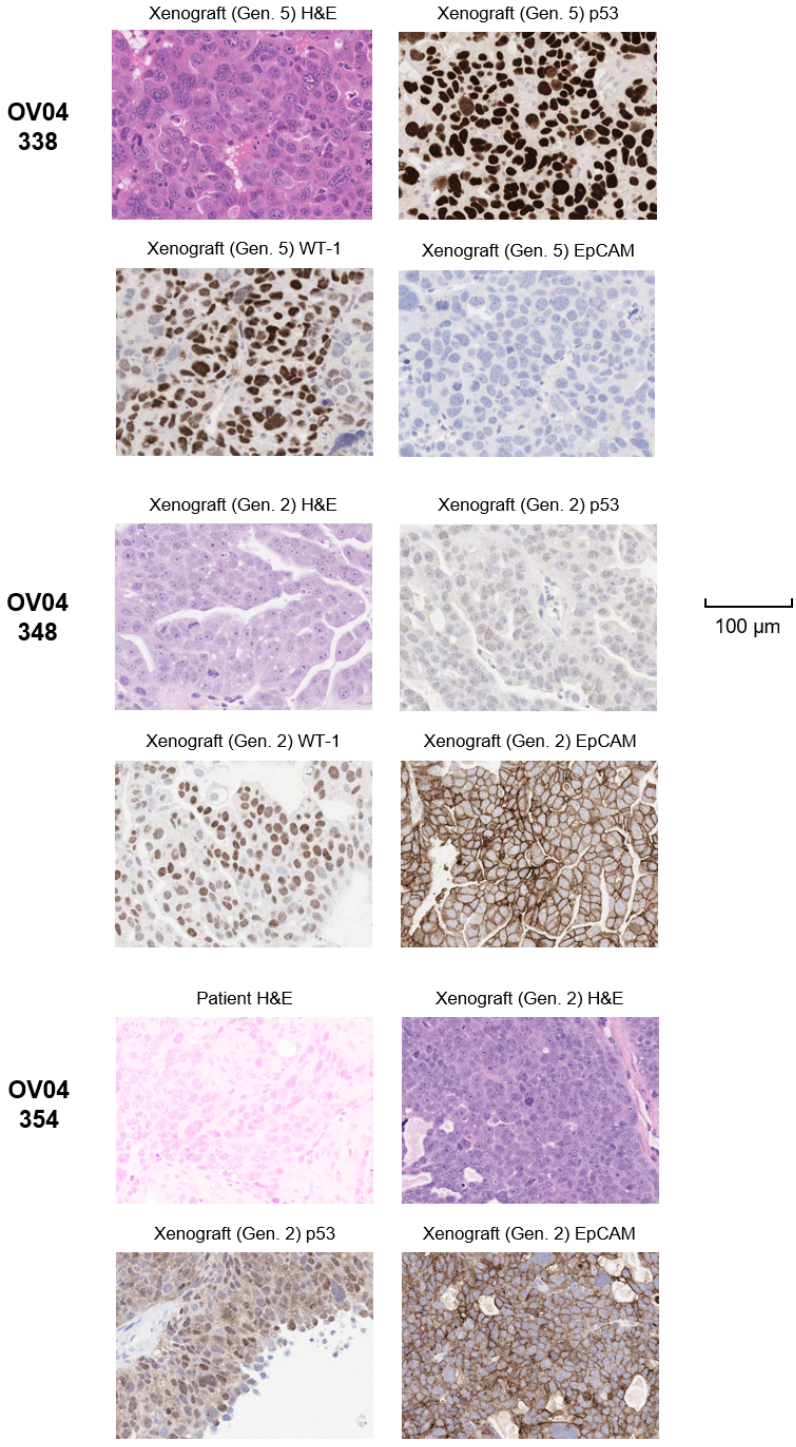


Figure 3.5: Validation of xenografts by histology (cont. 2/5)

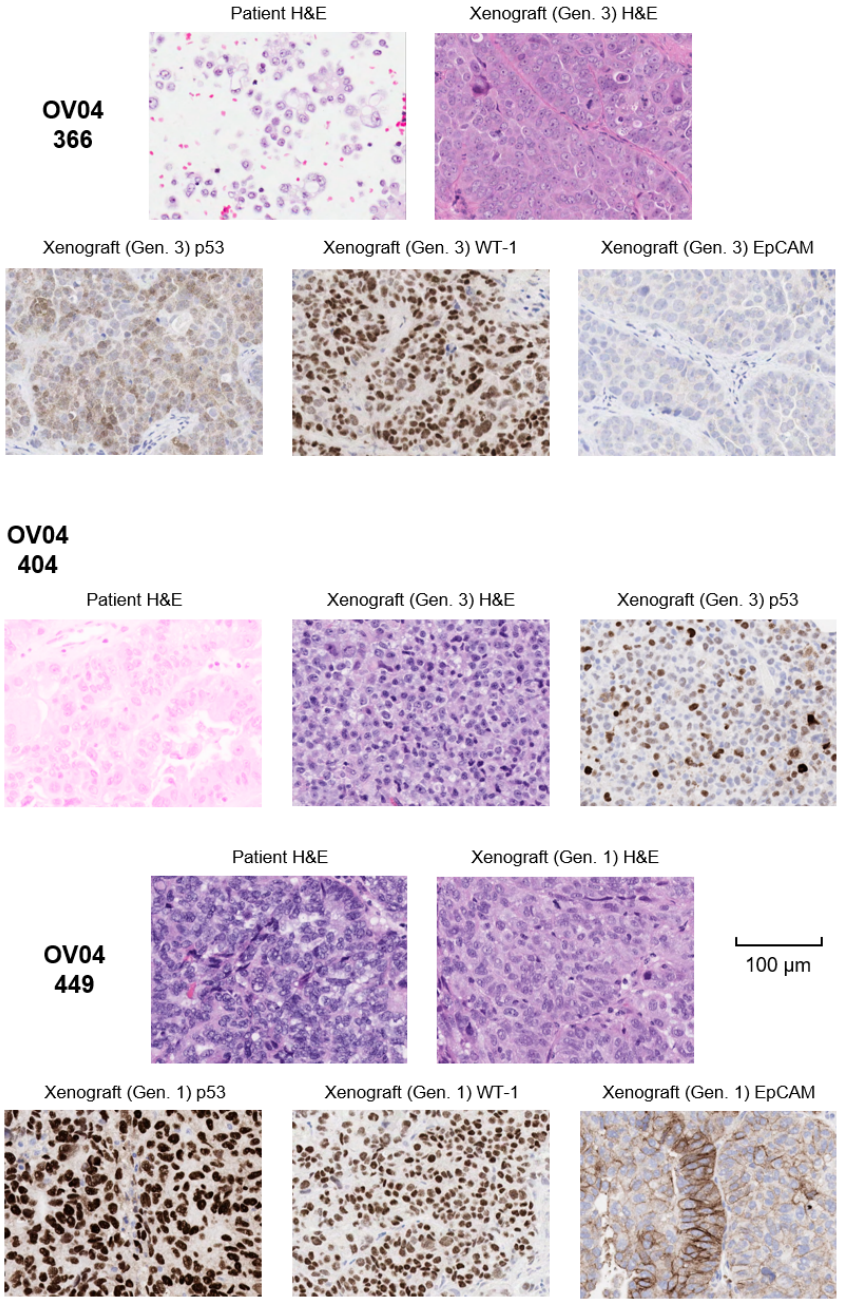


Figure 3.6: Validation of xenografts by histology (cont. 3/5). Note that OV04-404 was discarded based on the p53 IHC shown here which appears strongly wild-type.

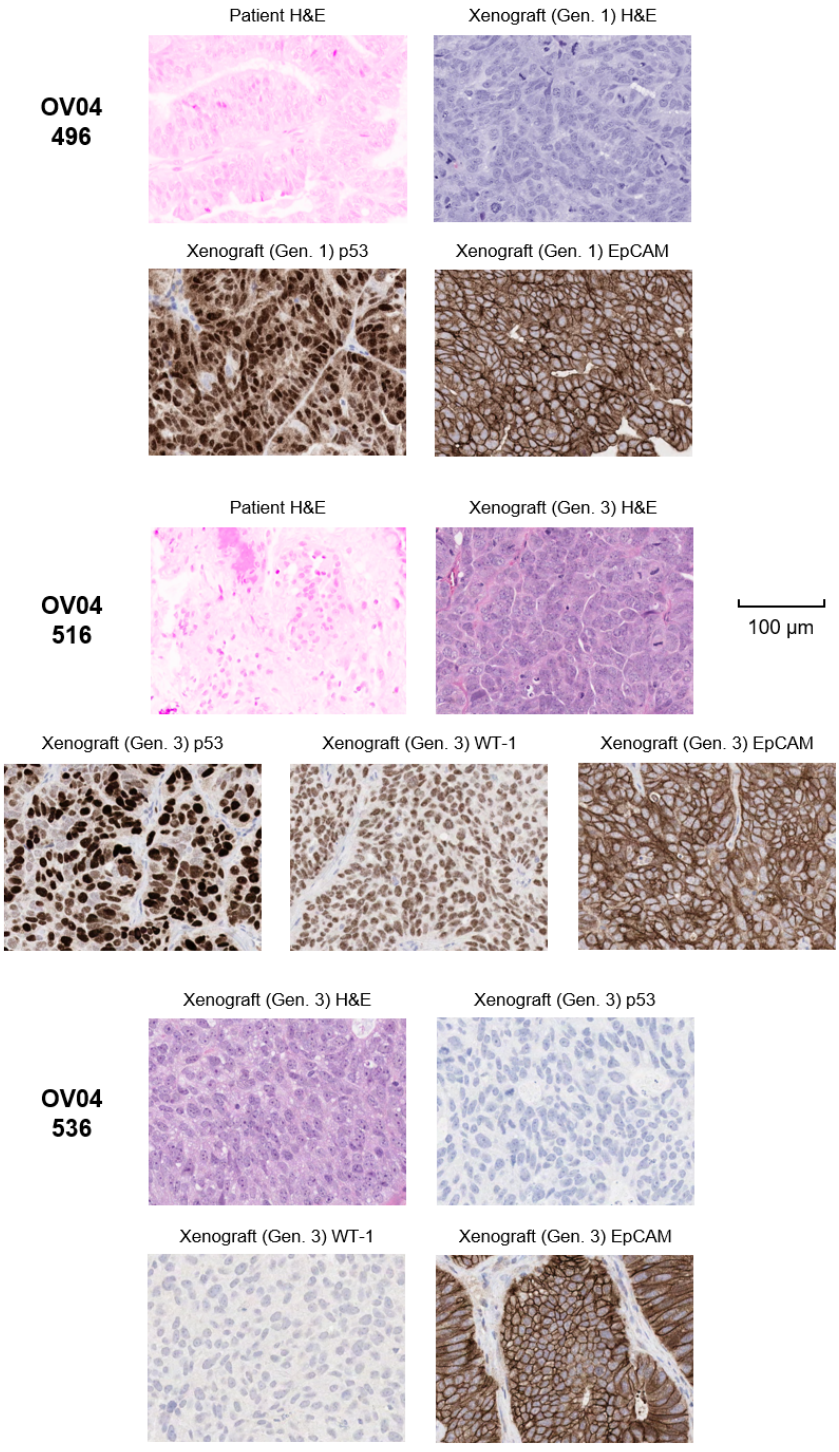


Figure 3.7: Validation of xenografts by histology (cont. 4/5)

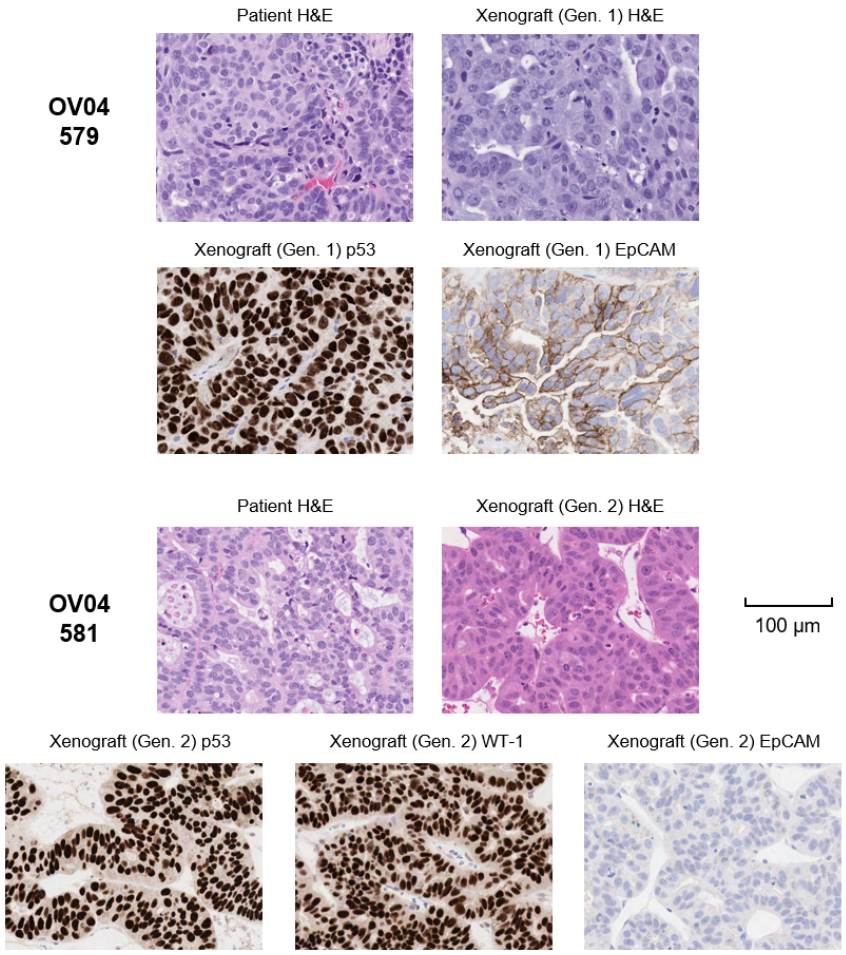


Figure 3.8: Validation of xenografts by histology (cont. 5/5)

3.4.3 Validation by Tagged Amplicon Sequencing

HGSOC is driven by ubiquitous early TP53 mutation. Therefore, targeted sequencing of key loci in xenografts and original patient tumours represents an efficient way to (a) confirm HGSOC status, (b) further validate that the xenograft is representative of the original sample, and (c) determine a little more detail about each sample. DNA was extracted from samples of both the PDX and the original tumour (partly by me, and partly by other members of the Laboratory as part of other projects), and put through Tagged Amplicon Sequencing (TAm-Seq) using the Brenton Lab's Optimal Panel (an improved version of the previously published Basic Panel (Forshew et al., 2012) by Maria Vias. TAm-Seq uses a panel of specific primers to perform targeted-sequencing across key regions of the genome, including the full length of *TP53*, along with other genes implicated in HGSOC. These include:

- *BRCA1* and *BRCA2*
- Genes coding for proteins associated with BRCA (*BARD1*, *BRIP1*, and *PALB1*)
- A number of genes associated with homologous recombination (*RAD51B*, *RAD51C*, *RAD51D*, *FANCM*)
- ATK pathway related genes (*PTEN*, *EGFR*)
- Checkpoint related genes (*CDK12*, *RB1*)
- The RAS inhibitor *NF1*

The panel also captures key hot-spots associated with other ovarian subtypes: *BRAF*, *KRAS*, *PIK3CA*, *NRAS*, *CTNNB1* (although *PIK3CA* mutation is seen at low level in HGSOC cases). While *TP53* mutation is seen ubiquitously in HGSOC, the range of mutations seen is extremely wide. Therefore, the exact *TP53* mutation can act as a signature for each tumour, and allow confirmation of the presence of the same initial driver event.

The TAm-Seq data for *TP53* is shown in Table 3.3. OV04-536 lacks any *TP53* mutation. This suggests that, despite crude histology consistent with HGSOC and the complete abolition of p53 staining (as detected by IHC), this sample may not be HGSOC. This sample was clinically assessed as mixed HGSOC/endometrioid, meaning the origin of the xenograft is not clear. Given this information, one of two possibilities seem likely: either the resulting xenograft is endometrioid but with certain similarities to HGSOC, or it is an HGSOC-type tumour where p53 function is lost as a result of mutation in a related protein. Given the cryptic and potentially unrepresentative nature of this sample, it was excluded from future experiments as a precaution.

Xenograft validation by TAm-Seq			
PDX	PDX *TP53* mut.	Patient	Patient tumour *TP53* mut.
OV04-072 xenograft (G1)	c.574C>T	OV04-072	c.574C>T
OV04-075 xenograft (G2)	c.991C>T	OV04-075	c.991C>T
OV04-297 xenograft (G1/3/4)	c.661G>T	OV04-297	c.661G>T
OV04-338 xenograft (G1)	c.637C>T	OV04-338	c.637C>T
OV04-348 xenograft (G1)	c.637C>T	OV04-348	c.637C>T
OV04-354 xenograft (G2)	c.759delC	OV04-354	c.759delC
OV04-366 xenograft (G2)	c.783delT	OV04-366	c.783delT
OV04-449 xenograft (G2)	c.403T>G	OV04-449	c.403T>G
OV04-496 xenograft (G1)	c.1024C>T	OV04-496	c.1024C>T
OV04-516 xenograft (G1)	c.742C>T	OV04-516	c.742C>T
OV04-536 xenograft (G1)	None	OV04-536	None
OV04-579 xenograft (G1)	c.536A>G	OV04-579	c.536A>G
OV04-581 xenograft (G1/2)	c.743G>A	OV04-581	c.743G>A

Table 3.3: Summary of the results of TAm-Seq performed on the xenograft samples used for scRNA-Seq. The fingerprint of a single key mutation in p53 is conserved between these samples and the patient samples from which they are derived. Note the sample without a mutation, which is likely endometrioid rather than HGSOC. Numbers after xenograft ID indicate the generation of PDX tested, with G1 indicating mice receiving tumour material direct from the patient, G2 recieving material from G1, etc. The ('generation' nomenclature is used rather then 'passage' as the latter is unclear, sometimes starting from the first transfer from human to mouse, and sometimes only from the first transfer from mouse to mouse. Where multiple generations are listed, all were tested and gave similar results.

3.4.4 Validation by STR genotyping

To further confirm that all remaining PDXs are good recapitulations of the patient tumours from which they are derived, DNA was sent for STR genotyping by Genetica Laboratories. Analysis of the sequencing results was performed by the CRUK Cambridge Institute Biorepository team. Methodological details are given in subsection 2.6.1. A summary of the results is shown in Table 3.4. The general trend is of a similar STR fingerprint from both PDX and original tumour for each sample, showing that the PDXs broadly model the original patient disease. Where disparities are seen, in all but one case, two STR alleles are seen in the patient sample with only one being present in the PDX, suggesting that there is a greater degree of heterogeneity within the patient tumour than has been recapitulated in the xenograft. In a single case (D13S317 locus, OVO4-072) two alleles are seen in the patient sample and three in the xenograft; this is likely to represent outgrowth in the xenograft of a minor clone that is below detection threshold in the original patient sample. These variations are rare enough to support the conclusion that overall the xenografts are a good model of patient disease; however, they reinforce the data from Siru's work that show that transplantation experiments (be they *in vivo* or *in vitro*) are stressful events in which the vast majority of cells are lost, and the tumour engrafts from only a small fraction of surviving cells. It is thus important to remain mindful that creating models of any kind from primary tumour is, under even the best conditions, going to result in a large degree of selection. Any

List of passaged xenografts																		
OV04 No.	Type	Gen.	D3S1358	TH01	D21S11	D18S51	Penta E	D5S818	D13S317	D7S820	D16S539	CSF1PO	Penta D	vWA	D8S1179	TPOX	FGA	AMEL
449	tumour		14	6/9	28/29	12/14	11	11	9	9/11	12/14	9/12	13/16	14	16	8	20/21	X
449	xeno	2	14	6/9	28/29	12/14	11	11	9	9/11	12/14	9/12	13/16	14	16	8	20/21	X
297	ascites		17/18	**6/9**	27/33.2	13/15	12/18	11/12	11	9/10	9	13	8/12	14/17	10	8/11	22	X
297	xeno	1	17/18	**9**	27/33.2	13/15	12/18	11/12	11	9/10	9	13	8/12	14/17	10	8/11	22	X
297	xeno	4	17/18	**9**	27/33.2	13/15	12/18	11/12	11	9/10	9	13	8/12	14/17	10	8/11	22	X
338	ascites		16	9.3	28/30	14	10/11	13	11	8/9	11	10	13	16/18	10/16	8/11	22	X
338	xeno	1	16	9.3	28/30	14	10/11	13	11	8/9	11	10	13	16/18	10/16	8/11	22	X
348	ascites		14/18	6/9	30	19	5	11/12	8/11	8/10	11/12	10/13	12/14	16/19	8/10	11	24	X
348	xeno	1	14/18	6/9	30	19	5	11/12	8/11	8/10	11/12	10/13	12/13	16/19	8/10	11	24	X
366	ascites		14	9.3	30/31	10/17	7	12	12	10/12	12	10	11	16	11/15	8/11	21	X
366	xeno	2	14	9.3	30/31	10/17	7	12	12	10/12	12	10	11	16	11/15	8/11	21	X
409	ascites		18	8/9	28/30	10/18	15/16	11	12	9/14	11	10/11	11	17/18	14/15	8	22/23	X
409	xeno	2	18	8/9	28/30	10/18	15/16	11	12	9/14	11	10/11	11	17/18	14/15	8	22/23	X
581	tumour		17	6	30/31	15	12/14	12/13	12	9/11	**12/13**	10	13/15	18/19	13/14	11	20	X
581	xeno	1	17	6	30/31	15	12/14	12/13	12	9/11	**13**	10	13/15	18/19	13/14	11	20	X
581	xeno	2	17	6	30/31	15	12/14	12/13	12	9/11	**13**	10	13/15	18/19	13/14	11	20	X
496	tumour		15/17	7/9.3	30/32.2	11/13	11/16	11/12	**11/14**	8/10	9	10/11	10	15/17	13/15	8	21.2/25	X
496	xeno	1	15/17	7/9.3	30/32.2	11/13	11/16	11/12	**11**	8/10	9	10/11	10	15/17	13/15	8	21.2/25	X
516	tumour		14/15	**6/9.3**	29/32	**10/16**	12/16	**10/11**	11	10/13	**8/12**	10/12	11/13	14/19	12/13	8	**25/27**	X
516	xeno	1	14/15	**6**	29/32	**16**	12/16	**10**	11	10/13	**8**	10/12	11/13	14/19	12/13	8	**25**	X
579	tumour		14/16	6/9.3	28/30	12/16	7/10	12	8/14	9	12	12	10	18/19	13/14	8/11	**20/23**	X
579	xeno	1	14/16	6/9.3	28/30	12/16	7/10	12	8/14	9	12	12	10	18/19	13/14	8/11	**23**	X
354	tumour		**15/16**	6	28/30	14/15	5/12	11/13	**8/11**	8/9	**8/13**	11/12	**11/13**	16/17	**8/13**	8/12	19/24	X
354	xeno	2	**16**	6	28/30	14/15	5/12	11/13	**11**	8/9	**8**	11/12	**13**	16/17	**8**	8/12	19/24	X
072	ascites		17/18	**9/9.3**	**29/30**	**13/14**	12/17	10/11	**11/12**	8/9	13	10/13	**9/12**	14/17	12/16	9/11	**22/23**	X
072	xeno	1	17/18	**9**	**30**	**14**	12/17	10/11	**11/12/13**	8/9	13	10/13	**12**	14/17	12/16	9/11	**22**	X
075	ascites		16	7/8	**28/31.2**	18	10	11/12	9/11	9/11	**11/12**	10	9/13	17/19	10/12	8/11	23	X
075	xeno	2	16	7/8	**31.2**	18	10	11/12	9/11	9/11	**12**	10	9/13	17/19	10/12	8/11	23	X
294	ascites		16	9.3	29/30	12	7	11	11	7/8	11	11	9	17	11/13	9	22/23	X
294	xeno	1	16	9.3	29/30	12	7	11	11	7/8	11	11	9	17	11/13	9	22/23	X

Table 3.4: Summary of STR results. Table summarises the variants (signified by number of short tandem repeats in each) for the 16 loci within the Promega PowerPlex 16HS Cell Line panel, for each sample. Differences between patient and xenograft samples are highlighted in bold.

experimental design or analysis must be sure to take account of this limitation.

3.4.5 Validation by sWGS

As HGSOC is primarily a cancer caused by p53 mutation leading to aberrant genomic copy-number, STR is probably an effective method to confirm the origin of the material being used, but may not be the best method to detect complex shifts in genomic heterogeneity as a result of tumour evolution, since it is not able to detect duplications in large genomic fragments. In order to further confirm that genotype is maintained across the xenografting process, DNA from patient tumour and associated xenografts was put through shallow whole-genome-sequencing (sWGS), and relative copy-number changes across the genome profiled. This produced a characteristic 'fingerprint' for a tumour based on the specific pattern of duplication and/or deletions of sections of genome. This was performed by Maria Vias as part of another project in the Laboratory which utilised the xenografts developed during my work. These data are shown in Figure 3.9.

Similar patterns are seen in each paired sample, indicating similar patterns of genomic instability are present in both patient and xenograft tumours. The biggest variation observed is the changes in the magnitude of copy number change between samples (vertical 'stretching' of the entire plot). This variation is a methodological artefact rather than any real biological difference, and arises because it is hard to accurately calculate absolute copy number for a given portion of the genome. As such the algorithms behind such calculations work from the assumption that average copy number over the entire genome is not significantly non-diploid, and calculate a relative copy number for any given stretch of the genome by comparing against average copy number across the genome (and extrapolating absolute values from the assumption of diploid character). However in HGSOC this underlying assumption is not reliable, as many tumours have undergone significant (and potentially repeated) duplications of the genome. As such the absolute values (and the y-axis amplitude) shown are unreliable as absolute copy number is not being accurately calculated. However the characteristic shape of the graph can be seen to remain the same, even if it has been 'stretched' vertically, as the relative copy number across the genome will still be being accurately calculated.

This supports the assumption that xenografts are representative of their original tumours, not just at the STR level (which will be minimally affected by genomic instability), but also at the level of shallow Whole Genome Sequencing, which would be expected to more clearly show a significant loss of genomic heterogeneity. For example, the domination of a xenograft by outgrowth of a divergent minor clone would be expected to produce a somewhat different 'genomic fingerprint', since while the type of genomic fragmentation would likely remain the same (due to similar shared driver effects) the places where fragmentation occurred would be expected to gradually diverge over time between the different clonal populations. The results shown here suggest either that there is minimal loss of genomic heterogeneity (hence similar patterns of fragmentation), or that there was little to begin with. In either case, this would indicate that xenografts represent a good model of the original tumour.

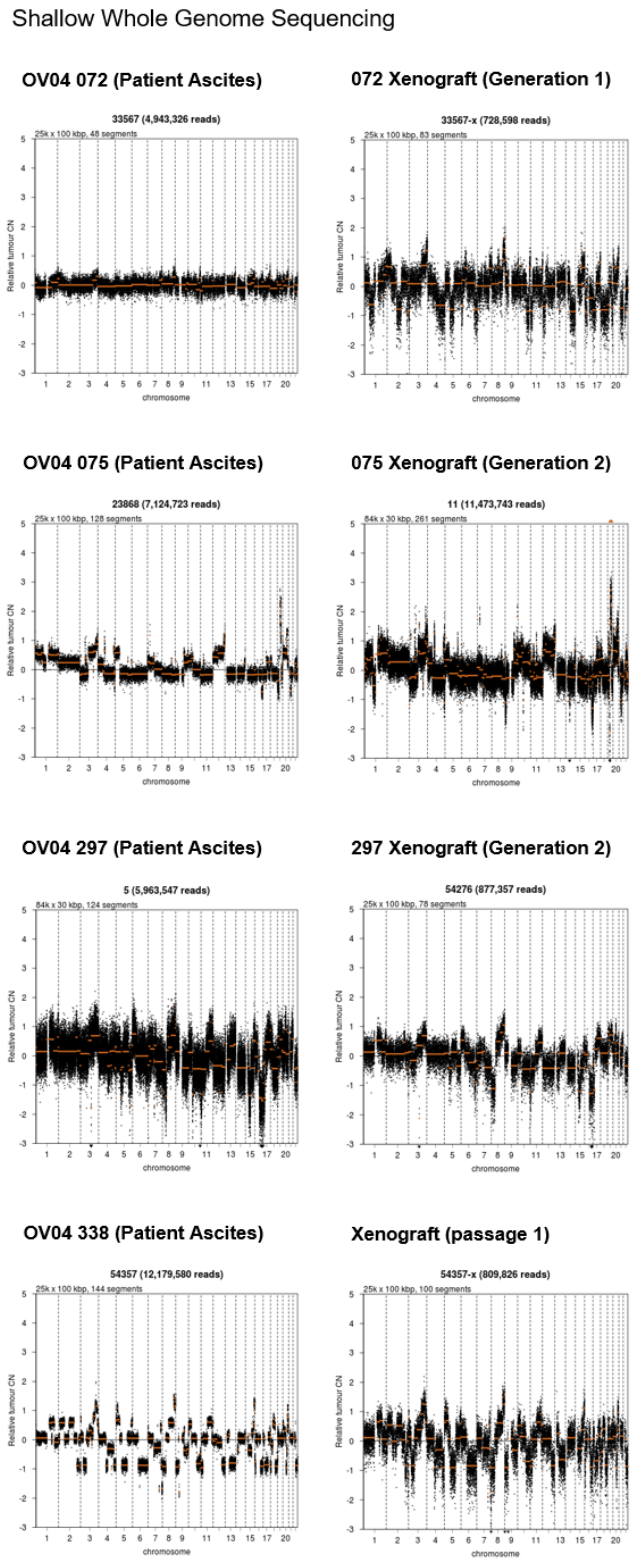


Figure 3.9: Validation of PDXs by Shallow Whole Genome Sequencing (1/3). The similar patterns for patient and PDX tumours indicate similar genomic disarray showing the PDXs are good recapitulations of their tumour of origin (as well as further confirming PDX HGSOC status).

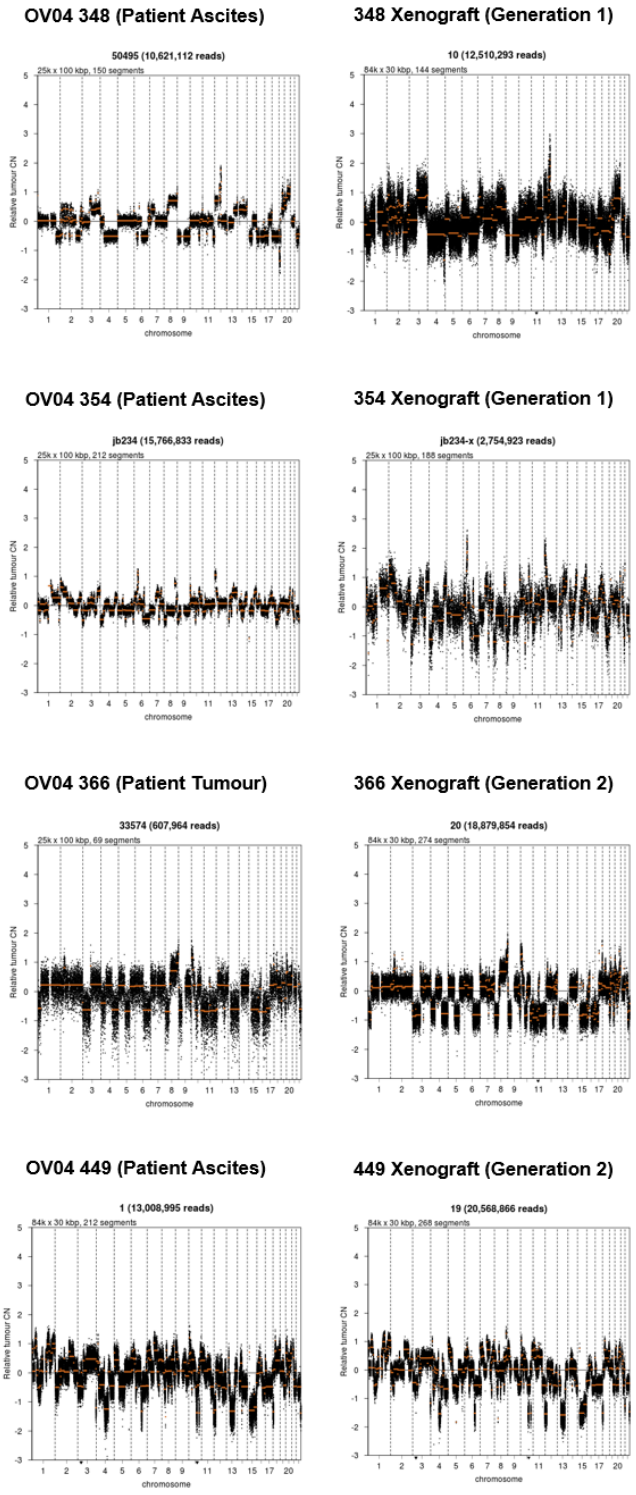


Figure 3.10: Validation of PDXs by Shallow Whole Genome Sequencing (2/3).

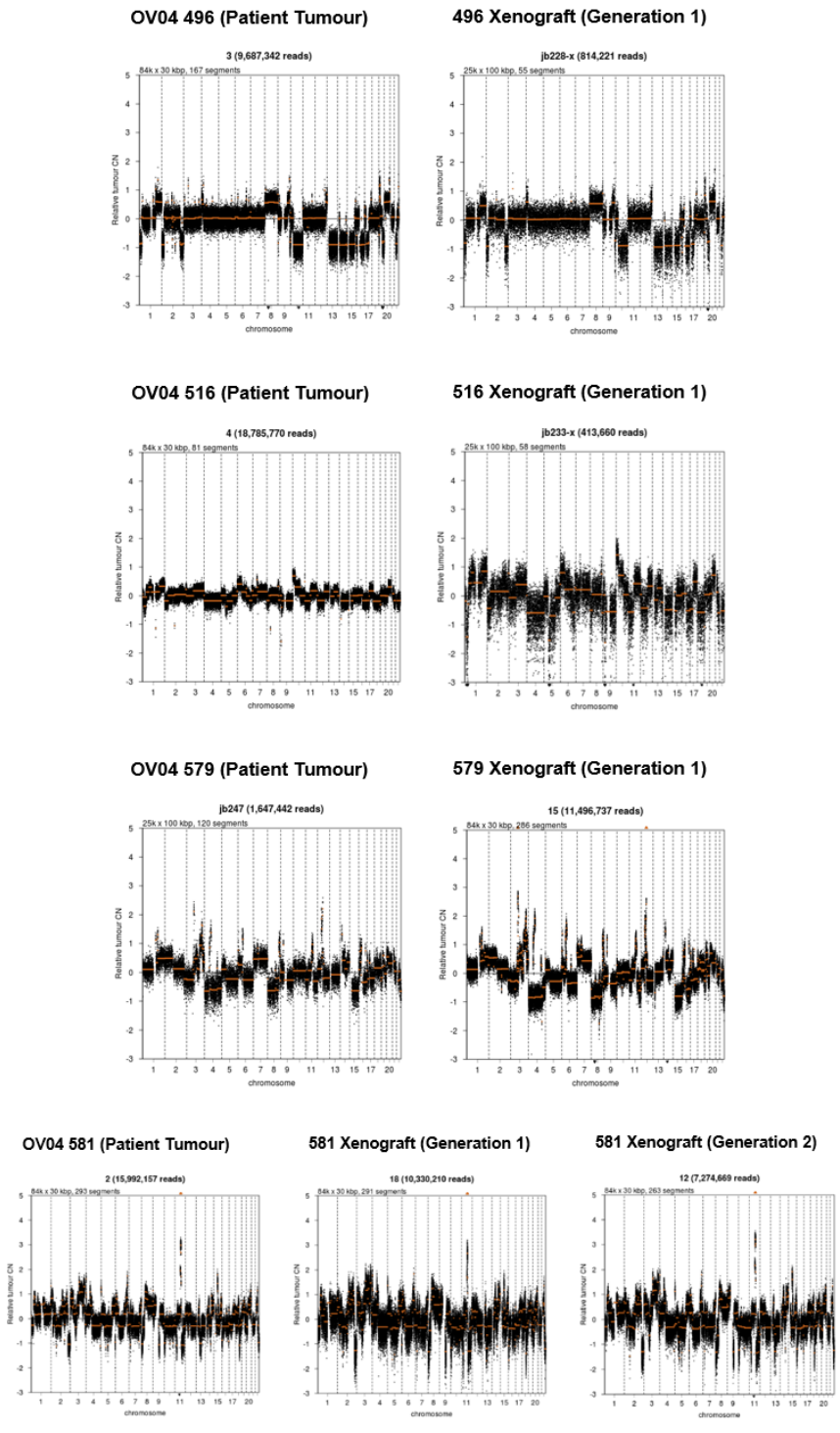


Figure 3.11: Validation of PDXs by Shallow Whole Genome Sequencing (3/3).

3.4.6 Frozen samples

Frozen dissociated cells were widely used during early efforts at implanting tumours, but as part of troubleshooting efforts, this was discontinued because of the risk that freezing was detrimental to viability and therefore may have played a role in the failure of early implantation work. Once successful passaging was achieved, small pieces of tissue were frozen down in freezing media to determine if passaging could be achieved from frozen stocks. Re-implantation from frozen PDX pieces proved highly efficient, with the same near-perfect technical success rate of fresh pieces. However, the lead-time between implantation and tumour growth noticeably increased when growing from frozen, likely a result of reduced cell viability and therefore regrowth from a smaller initial population. The use of frozen dissociated cell solution was not rigorously investigated further.

3.4.7 Assessment of mechanical digestion

Given the available data, it appears that surgical implantation is an effective method of establishing representative xenografts from patient material, and conducting serial passage. However, it is an involved and time-consuming procedure, with moderate ethical burden, and potential for refinement to reduce ethical impact and time expenditure. This would also allow for larger studies with the same manpower. One option, used extensively in the breast cancer field, is to mechanically dissociate tissue in order to generate a single-cell solution which may be injected. This is very hard with ovarian patient material which is typically extremely tough, and very fibrotic. However, serially passaged xenografts do not generate this fibrosis, and are much more manageable; thus a mechanical approach might improve efficiency of xenograft passaging. Based on a detailed protocol kindly provided by Peter Eirew (British Columbia Cancer Research Centre), a methodology was adapted for use with ovarian xenografts (see subsection 2.1.3), and tested for passaging of four xenograft lines, in parallel with equivalent tumours engrafted by the surgical method. All implantations (100 technical replicates) generated this way grew, suggesting that both methods are adequate for this purpose.

However, a couple of interesting observations were made during this experiment which could be important for future studies in this area. The first is that while technical replicates generated by this method still have similar lead-times within a cohort, the lead-times vary significantly between samples. For the samples with ascitic origin, the lead-time was reduced from $\sim$2.5 months to $\sim$2 months. For the samples of solid tumour origin, the lead-time increased from $\sim$3 months to 5-6, months. While it is difficult to draw strong conclusions from such a small pilot experiment, it is interesting that the inverse effect on lead-time is seen in the two types of samples which had similar lead-times when passaged by surgical implantation. The second observation emerged when this methodology was used to prepare single-cell solution for other experiments. A cell count was performed on the single-cell solution using a haemocytometer, at which point it became clear that only $\sim$1% of the

cells were still live, versus the 80%-90% that is typical for enzyme digested samples. The enrichment of dead cells seems to suggest that either that mechanical dissociation is highly toxic, or that it preferentially liberates dead cells. Both rationales could be plausibly argued; cells are liberated from tissue where the blade cuts, but these cells are heavily damaged in the process – or that because cells that are already dead would be expected to be less strongly attached, they would be more easily detached from the tissue by mechanical agitation.

In either case, the number of live cells present is sufficient for functional passage. However, the use of this method may have implications for selection and loss of diversity, and this limitation should be considered when deciding whether or not to use this method. It is unclear exactly how the effect on lead-time arises, but it appears (on the limited evidence here) to reflect the differing origins of these two types of cells. Ascites would by nature be expected to represent more advanced disease, able to thrive in isolation and in an unattached fashion, and able to form xenografts by injection while solid tumours are not; this likely reflects enrichment of a more metastatic phenotype. It is possible that the high attrition seen during the mechanical digestion process may impact the two types in different ways. For an ascitic sample, which is on average more resilient due to its origin, there may be a significant proportion of cells able to survive the process of implantation and proliferate, such that the stress of attrition simply acts as a selection pressure, enriching for the hardiest cells. For cells of solid tumour origin, where the ability to survive implantation is rare, the depletion due to the stress of mechanical digestion may simply cut down an already limiting pool of cells, resulting in a slower recovery and increased lag-time. If this is the case, then the way this aggressive metastatic nature is conserved over generations, despite Siru's work demonstrating that the prevalence of phenotypes fluctuates significantly between engraftment ($EpCAM^-$-driven) and recapitulated tumour (mixed and $EpCAM^+$ dominated), would seem to imply these characteristics are at least partly encoded at a level higher than that of phenotype. This fits well with our understanding of ascites cells as representing advanced disease predisposed to metastasis by accumulation of a higher degree of genomic aberration.

3.4.8 Validation of epithelial status

Some concerns have been raised about the immunohistochemistry (IHC) staining for EpCAM within this validation, noting that several samples do not stain strongly for EpCAM by IHC, and questioning whether these cells are in fact epithelial or indeed even human. That these cells are unambiguously human is demonstrated by aberrant p53 IHC, and the TAm-Seq results, as mouse cells would not show a characteristic pattern of p53-driven DNA damage consistent with HGSOC. This also strongly implies that these cells are epithelial, or at least that they are derived from an epithelial lineage. In the next chapter these cells are also shown to stain strongly with human antibodies, something which is not observed with mouse tissue (see Figure 4.7). Finally, it is worth noting that in many samples, clear pseudo-fallopian structures can be observed, which would not be expected from a non-epithelial murine tissue.

For example, OV04-581 (see bottom row, Figure 3.8) shows only weak EpCAM staining by IHC, but also clear fallopian-like structures (which can be seen particularly clearly as a result of the contrast in the p53 IHC staining in the bottom left of this figure).

To further confirm these cells are epithelial in nature, three PDX samples with weak EpCAM staining (OV04-338, -366, and -581) were stained for additional epithelial markers. EpCAM immunohistochemistry was repeated in parallel with IHC for cytokeratin-7 (CK7) and cytokeratin-18 (CK18), which are both epithelial markers known to be strongly expressed in fallopian and HGSOC tissue. The results are shown in Figure 3.12. Thirteen PDX samples were IHC stained with each cytokeratin. CK18 was strongly expressed in all samples and in all cells which appeared to be tumour by crude morphology. The vast majority of samples showed similar staining for CK7, however OV04-449 showed patchy staining for CK7, with $CK7^{weak}$ cells and $CK7^{strong}$ cells divided into discrete patches, all of which were morphologically epithelial, and showed strong staining for CK18. These observations are curious, but do not impact the identification of the cells as epithelial. These observations will be discussed in more detail in section 5.3.3.

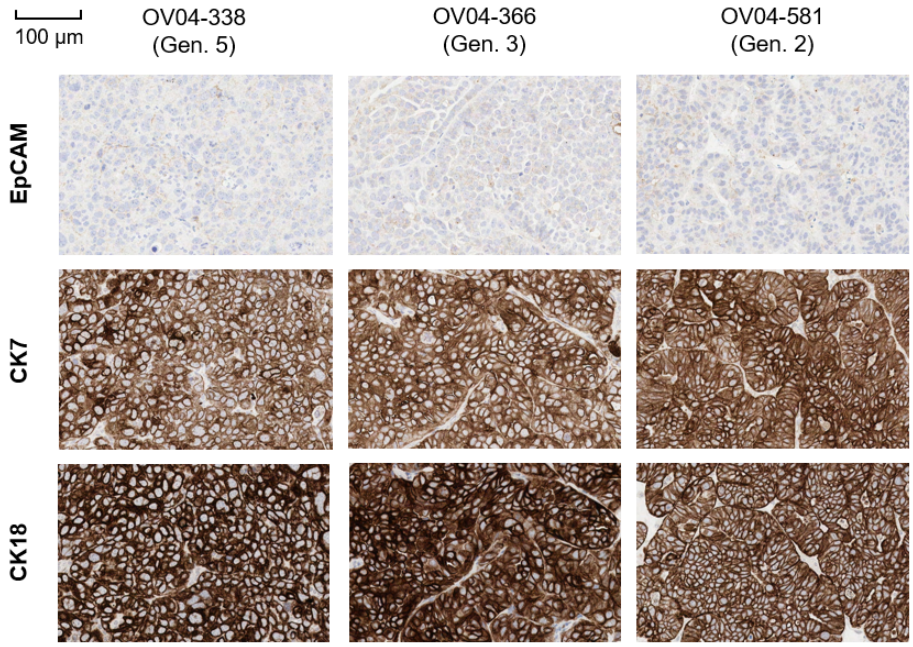


Figure 3.12: Immunohistochemistry staining for EpCAM, cytokeratin-7 (CK7), and cytokeratin 18 (CK18), for three PDX samples with low expression of EpCAM. PDX generation numbers are shown. Images are sections from the same tumour block, and while not perfectly matched, were taken from the same area of each section to maximise equivalence.

These results clearly demonstrate that even the cells with low EpCAM by IHC are clearly epithelial in nature, validating the status of these cells. It has been noted that expression

of EpCAM in these samples by IHC is far weaker than might be expected given the strong signals for EpCAM seen in these samples by flow cytometry in subsequent chapters, (e.g. in Figure 4.4). There are at least two possible explanations for this. The first is that a different antibody clone is used for the two techniques (HEA-125 for flow cytometry; VU-1D9 for IHC), so the difference may be due to varying sensitivity. This would require some significant difference in EpCAM at molecular level. It would require most samples to express a type of EpCAM that can be bound by both antibodies, but for a handful of samples to express a structurally modified variant of EpCAM (either at the protein level or post-translationally) to which only HEA-125 can bind.

An alternative explanation would be that the effect arises as a result of inherent differences in the methodologies of the two techniques. It is important to note here that the hypothesis underlying this thesis is not that EpCAM has a key functional (or aetiological) role in phenotypic heterogeneity in HGSOC, but rather that phenotypic heterogeneity exists within HGSOC, and that the (unknown) mechanisms which underlie this also result in differential expression of EpCAM which can be used as a readout for the phenotypic subpopulations. As such, it is entirely reasonable that the degree of variation in EpCAM expression between samples may be higher than the variation within a given sample.

Immunohistochemistry and flow-cytometry, particularly in the context used in this work, have somewhat different methodological approaches. In particular, the IHC techniques used here follow a standard workup which was developed for use on any and all human tissues, not fallopian tube or IHC specifically, and so the expression levels over which the methodology is sensitive may be too low for some specific samples but adequate for others. The flow-cytometry parameters in this work were specifically developed for HGSOC tissue, for the purposes of maximising the differential between phenotypic subpopulations. Furthermore, the intensity of the technique's signal is entirely dependent on a wide variety of parameters which impact the sensitivity of the cytometer's detectors. As the machine is used for a wide variety of other experiments, these must be reset each time it is used, and their sensitivity adjusted until an appropriate level of signal is seen. As such, a flow cytometry signal of 10^4 in one experiment is not necessarily equal to a signal of 10^4 in another, unless the two were conducted in parallel. Due to the complexity of the following experiments, and the need to be parsimonious with mice, it was not common that different biological samples were assayed in parallel by flow cytometry.

As a hypothetical example, consider two biological samples A and B. Each has $EpCAM^+$ and $EpCAM^-$ subpopulations. For sample A, these have intensities of 10^3 and 10^2 respectively (in an arbitrary absolute scale), while for sample B they are 10^4 and 10^3 respectively. Given a cut-off of $10^{3.5}$ for detection sensitivity for IHC, sample A shows no EpCAM by IHC, while sample B is broadly positive (since $EpCAM^+$ are typically in the majority). When assessed in (separate) flow cytometry experiments however, both samples show an identical detection pattern, with an order-of-magnitude difference between the $EpCAM^+$ and $EpCAM^-$ subpopulations.

In its favour, this explanation does not require the mechanistic differences of the previous explanation. Against this explanation is that no prodigious inter-sample variation (i.e. more than one order-of-magnitude) was seen in any of the experiments in which samples were assessed by cytometry in parallel. However, the example above requires only a single order-of-magnitude difference, and with other factors impacting detection sensitivity (e.g. variation in cell numbers) it is entirely such small variations would have gone unnoticed.

Which of the above explanations is true is unclear (or indeed if there is an alternative not considered here). However, the key point from the above hypotheses is that different results from different techniques is not indicative of a logical deficiency in the work in question, but rather may be a result of the differing methodological intents of the two techniques, and their respective and varying technical limitations. It is even theoretically possible that one of these antibodies has an obscure off-target sensitivity that has not previously been observed. As previously outlined, since EpCAM is being used only as a marker, rather than in any mechanistic sense, even this unlikely possibility would not fundamentally undermine this work since the specificity of the antibody remains a valid discriminant of distinct phenotypic subpopulations.

3.5 Discussion

This objective of this method-development section of the project was to develop a suitable model for further study of the complexities of phenotypic heterogeneity in HGSOC. A PDX approach represents the gold-standard for a model of HGSOC capable of sustained growth in laboratory conditions. A variety of methods for generating PDXs from patient samples have been tested in an effort to improve upon a previously effective method (subcutaneous injection) which does not appear to be successful under current conditions. Based on systematic elimination of possible variables, the likely cause of this appears to be variability in the tumour samples being implanted. Several methods of immediate assessment and quality control (tumour imprinting; frozen sectioning) have been implemented to triage available samples, and ensure resources are targeted at only the most promising. Testing of alternative implantation methods (subcutaneous surgical implantation of tumour pieces) suggests this increases the successful implantation rate. These refinements have resulted in improved yields from these valuable tumour samples, and have also reduced the cost and ethical burden from unsuccessful implantations, in line with our ethical mandate to reduce the number of animals such work requires. These improvements have also resulted in a method which is simply more robust than its predecessor, allowing a wider, and thus more representative, range of PDXs to be developed.

Assessment of primary xenografts suggests that the relatively poor technical success rate, and highly variable lag-time in the growth of these samples, makes them unsuitable for practical experimentation. Furthermore, possible variability in their composition (due to heterogeneity in cellularity of patient material, and the nature of implanting solid tumour

pieces) may introduce variables which may compromise controlled experiments. In order to minimise these problems, the use of serially passaged xenografts was tested. Non-first-generation PDX samples demonstrate a near-100% success rate, and when implanted from the same original tumour, show a minimum of variability in growth rates and character. Thus the use of passaged xenografts allows both of the outlined challenges to be overcome. While these samples represent a model one step further removed from the original tumour, the fact they make experiments feasible makes this a necessary compromise.

A range of PDX samples have been successfully developed for subsequent experimental use. These samples have been validated using histopathological methods (H&E, EpCAM, p53, WT-1) to ensure they are HGSOC. To address the concerns discussed about the use of passaged xenografts, different passages of xenograft were assessed with STR profiling and sWGS. The results show excellent maintenance of the genomic characteristics of the original tumour over multiple passages, mitigating these potential concerns. With this evidence that genomic character is maintained, and Siru's data suggesting established tumours recapitulate original tumour phenotypic heterogeneity, there is strong evidence to suggest these xenograft models represent a high-quality experimental model for further work.

The PDX models developed represent samples from a mixture of backgrounds (as per Table 3.2), ensuring a representative mix of samples is available. A certain degree of selection is to be expected, since more aggressive samples will more readily engraft. It is worth considering that ascites cells presumably represent the vanguard of tumour cells – the most highly advanced disease able to thrive in solution and with fewer cell-cell contacts and niche/stromal effects. As such, while they may be efficient for generating models, they may not present the most representative picture of the disease as a whole. Therefore, it is important that experiments conducted in ascites can be demonstrably replicated in solid tumours. The establishment of models from both here ensure that both are available for downstream study.

These PDX models represent gold-standard HGSOC experimental systems, along with an established pipeline for the generation of future models from specific HGSOC tumours of interest (in particular from untreated HGSOC patients whose tumour samples the Brenton Laboratory now has access to and is developing PDXs from). These models do not represent a drastic step-change in the field, however they (along with the work of other laboratories) represent a 'latest-generation' model of disease. In particular these models are designed to be pragmatic solutions, eschewing complex orthotopic approaches in favour of a subcutaneous approach making the model in circumstances where resources, manpower, and advanced surgical skills are not available. In the context of this work the model also provides a 'home grown' line of PDXs meaning that the original patient material is available for comparative assessment; this is often a major limitation of dependence on an external model, where orginal material no longer exists (or is too limiting to be made available) meaning equivalence cannot be confirmed.

Chapter 4

Assessment and validation of xenograft phenotypic subpopulations

4.1 Introduction

The work discussed in the previous chapter covered validation of xenograft models to confirm they are HGSOC, and are representative of their tumours of origin. However, before meaningful characterisation of phenotypic subpopulations could begin, the presence and proportions of the subpopulations also needed to be validated to ensure that representative phenotypic subpopulations are being seen in the xenograft context. In addition, it was also important to ensure that the HGSOC human tumour cells could be efficiently separated from host mouse cells. A major challenge in Siru's work was the potential for the $EpCAM^-$ population to be contaminated with non-tumour cells, and the challenge of rigorously removing these. During the xenografting process, the amount of stroma transferred will be small, and unlike the tumour, this material is unlikely to significantly proliferate. As the HGSOC cells proliferate, the stromal and haematopoietic cells surrounding the tumour will be increasingly replaced by mouse equivalents. This provides additional options for filtering out these cells, based on species rather than lineage alone.

4.2 Methods

Methods for the dissociation of patient material were optimised during Siru's project, and were modified only slightly over the course of this project. However, she performed only a few

studies *in vivo*, and most of these did not require the subsequent dissociation of xenograft. As such, the existing protocol for patient tissue had to be adapted to the xenograft context. The final protocol is described in subsection 2.1.2 and can be contrasted with the original patient tumour dissociation protocol detailed in subsection 2.1.1. The biggest change is in the reduction of the length of the collagenase/hyaluronidase incubation (from 16 h to 2 h) since there is far less extracellular matrix in the xenograft tissues, compared to the extensive fibrosis and connective matrix seen in patient samples. Other changes, including the elimination of the ammonium chloride step, were made because the composition of the xenografts (e.g. minimal vascularisation) made them redundant, and these steps were eliminated in order to reduce the length of the protocol. The resulting single-cell solution was assessed by flow cytometry.

4.3 Results

4.3.1 Detection of flow markers using flow cytometry

During Siru's work, she defined the Tumour Initiating Cell population to be $EpCAM^{+}/PDPN^{-}/CD34^{-}$. This provided two good markers for separating out the Tumour Initiating Cell population from the rest of the tumour; CD34 is not very relevant in the xenograft context as it is primarily a stromal marker, and was used during Siru's work to eliminate non-tumour stromal cells. In the case of xenografts, the human stroma will have been outgrown by mouse stroma and this will be removed along with other mouse cells. Upon conducting pilot experiments, a problem immediately emerged. Xenografts were assessed by flow cytometry, and showed the expected heterogeneous mixed population of $EpCAM^{-}$ and $EpCAM^{+}$ cells. However, there was no heterogeneity for podoplanin, which appeared to be either entirely absent or so weakly expressed as to be effectively useless as a functional marker (see Figure 4.1 and Figure 4.2).

After confirming that this was not a sample-specific issue, and affected all xenografts, alternative anti-podoplanin antibodies were tested to determine if the specific clone being used was failing to detect the protein in the murine environment. Changing the antibody had no effect on the problem, suggesting that it was unlikely to be a technical issue, and that alternative options were needed. Looking back to the initial marker-screening work done by Siru, podoplanin was one of several markers with similar specificity that were identified in the screen, including CD90 (*THY1*) and CD10. She selected podoplanin as providing the best population separation, and discarded the other markers since they provided no useful additional data. In light of this, it was a logical next step to test CD90 and CD10 against the xenograft samples in case detection of these markers remained unaffected in xenografts. Both markers proved to be somewhat effective in resolving a positive and negative population, especially when also co-resolved with EpCAM. Both markers separated similar populations (replicating Siru's data) and CD90 was taken forward as it gave marginally better resolution

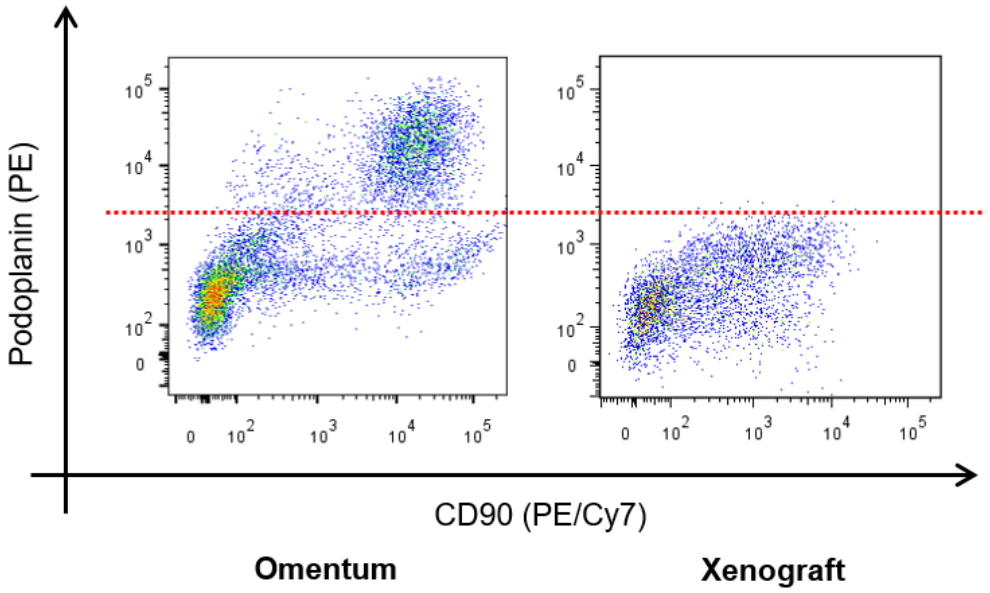


Figure 4.1: Podoplanin is expressed in original patient tumours, but in xenografts shows no elevation over control samples (more detailed xenograft sample/controls shown in Figure 4.2)

of the populations. To confirm that the CD90$^+$ subpopulation being separated is actually the same subpopulation as the podoplanin$^+$ fraction, patient tumours were subdivided based on EpCAM and CD90 status, and the podoplanin status of each of these assessed (see Figure 4.3). The results show that both CD90$^-$ subpopulations have no meaningful podoplanin staining, while the CD90$^+$ subpopulation is predominantly podoplanin$^+$. While CD90 may not entirely specifically isolate all podoplanin positive cells, it does appear to have near perfect sensitivity for podoplanin-positive cells (in that the CD90$^+$ fraction appears to contain all of the PDPN$^+$ cells, along with some PDPN$^-$ too). It is also worth noting that human stroma is known to contain cells positive for CD90 (Huynh et al., 2015) (Liu et al., 2004), so the CD90$^+$/PDPN$^-$ cells may not even be of tumour origin. If this is the case, CD90 would be an even better surrogate for podoplanin in xenografts where there will be no contaminating human stroma.

These experiments suggest that CD90 is a reliable surrogate for podoplanin. However, plotting the distributions of EpCAM vs. CD90 for both patient samples and xenografts raises interesting questions due to differences in the patterns observed (see Figure 4.4). Every HGSOC xenograft studied to date shows some variation on the same pattern – a diagonal spread ranging from a double-negative subpopulation up to a strong EpCAM positive subpopulation with some degree of CD90 positivity. In some cases, there is little CD90 staining and the diagonal tends to the vertical; in others, the EpCAM positive subpopulation begins to resolve into a CD90$^+$ and CD90$^-$ (e.g. OV04-338 and OV04-366, top and bottom-right of Figure 4.4). However, none of the xenografts show any subpopulation under (to the bottom-right of) this diagonal. In contrast, patient samples typically show a

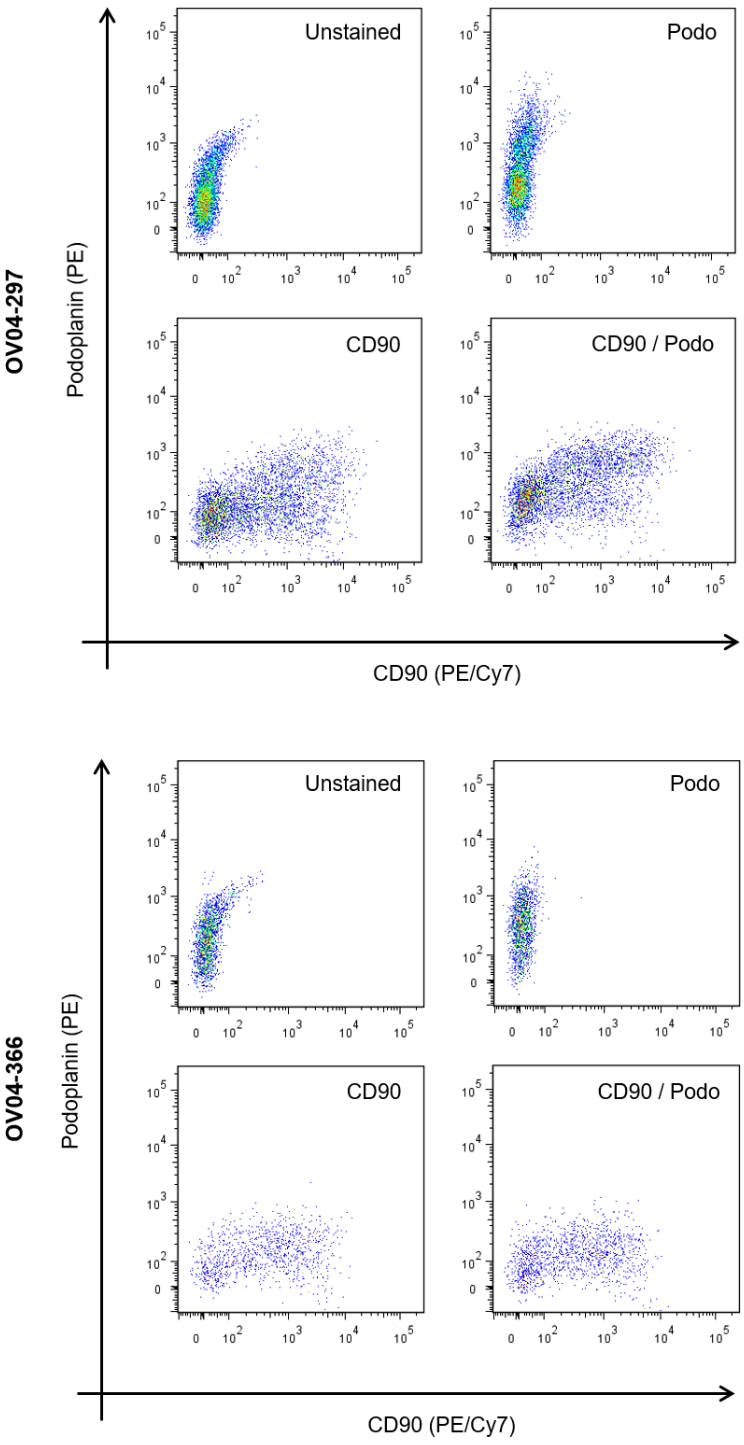


Figure 4.2: Full controls from the xenograft data shown in Figure 4.1 and a second sample. There is negligible difference between the unstained sample and that stained for podoplanin only (top row of each). When CD90 is stained for, there is negligible difference with or without podoplanin (bottom row of each).

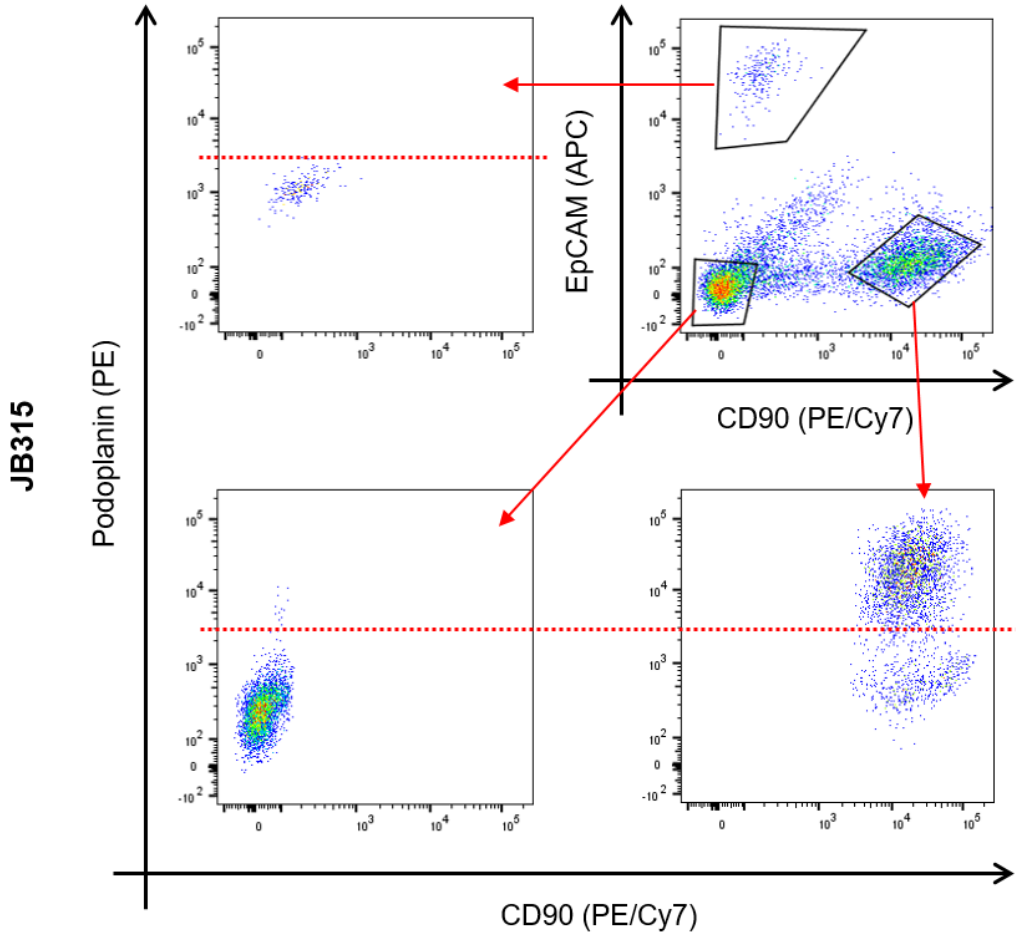


Figure 4.3: A patient tumour sample (top right) is split into subpopulations based on expression of EpCAM and CD90. The three subpopulations resolved this way are separated and their podoplanin expression shown. The $CD90^+/EpCAM^-$ population is the only one to show significant podoplanin expression (83% $PDPN^+$). This suggests that CD90 can reliably select out $PDPN^+$ cells, although it may potentially select some other cells too.

small EpCAM$^+$ subpopulation, a dense double-negative subpopulation (due to the degree of non-tumour contamination found in them), and an EpCAM$^-$/CD90$^+$ subpopulation not seen in xenografts. Given the previous findings that CD90 is a good equivalent marker for podoplanin positivity, it would be reasonable to conclude that this represents a genuine difference between the two types of sample rather than a technical deficiency in detection. Potentially the presence of this subpopulation in patient samples could be the result of contamination by stromal cells in the patient tumour (these cells would not express EpCAM, and stroma is known to contain podoplanin (Yamanashi et al., 2009) (Schoppmann et al., 2012) (Kawase et al., 2008) (Hoshino et al., 2011), however it is hard to correlate this with Siru's previous findings that this is the only subpopulation capable of tumour initiation.

These data raise many questions, and fully addressing these is somewhat beyond the scope of this project. Even within the scope of the aims of this chapter, the challenges of separating tumour and non-tumour cells in patient material make it hard to accurately determine the composition of pure patient tumour, and thus comparison of the subsequent xenografts to the material of origin remains challenging. The depletion of tumour cells as a result of chemotherapy also means the relative cellularity versus fibrosis and stromal cells may be quite low. As such, it may be that the PDX samples represent a'purer' model than their tissue of origin. To further complicate matters, the intent of this project is to generate a model akin to the established disease prior to treatment – since patient tumours undergoing surgery are almost invariably heavily chemotherapy treated, the makeup of the material of origin for the PDXs may be quite dissimilar to the subsequently established PDX due to the selective effects of chemotherapy. On this note it is interesting to observe that for samples of ascites origin, the patient and PDX profiles are often much more alike than is seen for solid tumour samples. This makes intuitive sense since live stromal cells would not be expected to be found free-floating in the peritoneum (at least not in a viable state), and so these samples would be expected to have a high tumour cellularity and minimal contamination (although variation due to chemotherapeutic selection would still be expected). This represents another reason why study of ascites samples is particularly valuable, bearing in mind previously stated caveats regarding ensuring any conclusions remain valid across the larger set of HGSOC tumours.

4.3.2 Purification of tumour cells from mouse host

The presence of stromal contamination is a major issue in the assessment of heterogeneity within patient tumour samples. The same is true of xenografts, however the xenografting process offers additional options for purifying out tumour cells from residual non-tumour, since in the xenograft context these cells will be of murine origin. Growing xenografts subcutaneously significantly reduces the degree of contamination since there is little natural stroma at the implantation site, and the tumour is able to recruit a limited amount of stroma, vasculature and other supporting host tissues to this site. This reduces the absolute number of contaminating cells that need to be removed, in contrast to patient samples

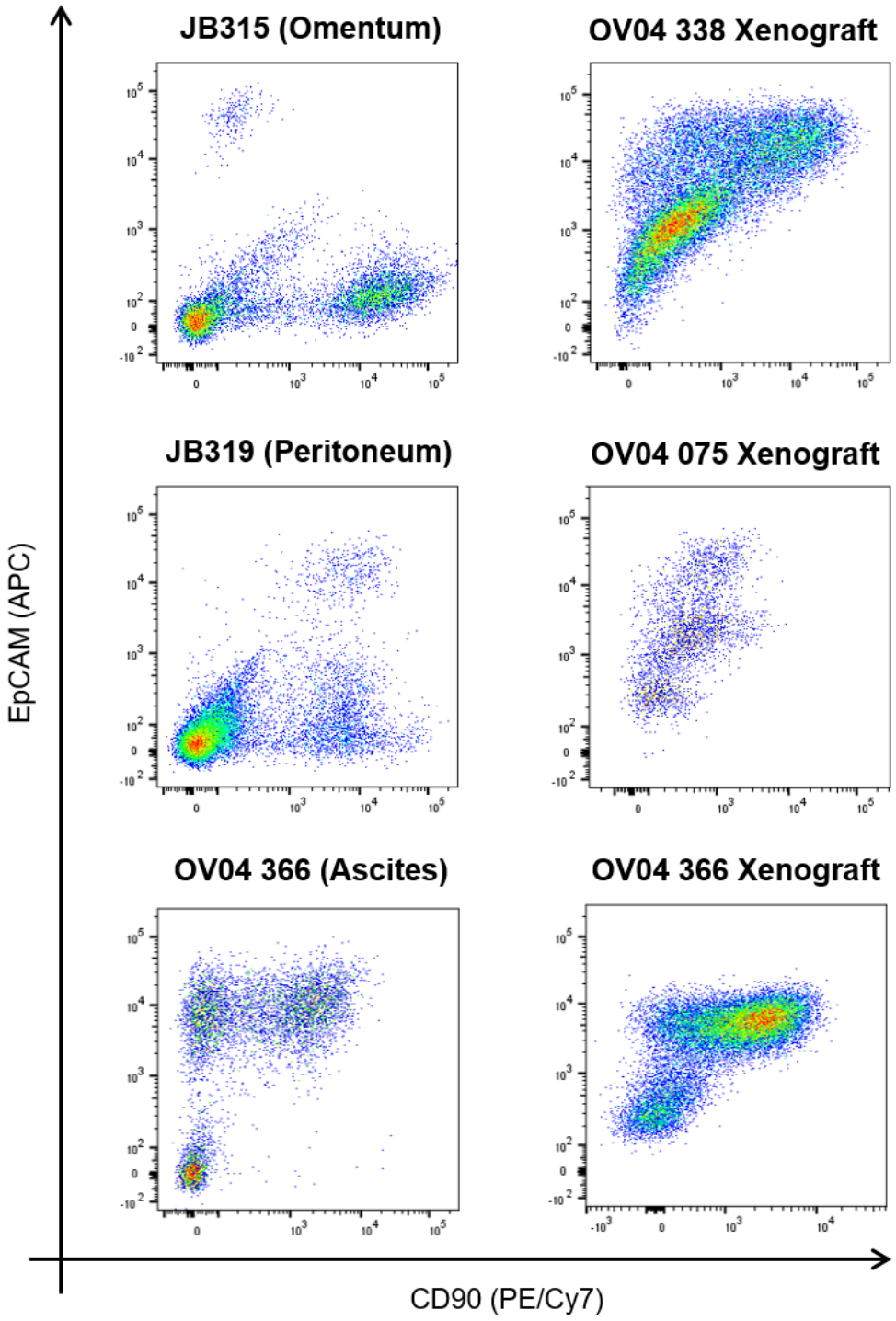


Figure 4.4: Sample distributions of EpCAM vs. CD90 for three patient samples (left) and three xenografts (right). The bottom two are a paired comparison between xenograft and inital sample used to generate it; note that this final patient sample is both of acites origin, and was frozen prior to assessment, while the top and middle ones were fresh solid tumour samples, which may explain the somewhat different distributions.

where the tumour is typically enmeshed in vasculature, stroma, untransformed epithelium, and fibrosis and other debris. Additionally, by growing the tumour in a different host species, non-tumour can be removed on the basis of species specificity, since even though non-tumour human cells will be found in the implanted sample, they will not proliferate in the mouse host, and this stroma/vasculature/etc. will be gradually be replaced by mouse tissue over the course of natural attrition. Thus in an established tumour, the only human cells should be tumour. Finally since the antibodies being used to identify subpopulations target human proteins, mouse cells should show little or no positivity for any of the markers being used, meaning contamination should remain clustered near the origin.

One approach for removing stroma in such situations is to deplete mouse cells using antibodies against relevant mouse lineages. A set of anti-mouse antibodies previously used in the Stingl Laboratory (CD31, CD45, Ter-119, and BP-1; see Table 2.1) were used (Eirew et al., 2008), along with an antibody targeting murine MHC-I, in order to deplete contaminating mouse cells. These antibodies were biotin conjugated, and detected with a streptavidin-APC/Cy7 secondary antibody. In order to test the effectiveness of this antibody cocktail, it was applied to several samples of mouse tissue which were digested in the same fashion as xenografts. These included bone marrow and heart (blood rich tissues to assess removal of haematopoietic cells), and subcutaneous skin tissues from a mouse flank with no tumour implanted. Subcutaneous fatty material was also tested but did not contain enough viable cells for meaningful assessment. As there were no human cells present, it would be expected that this mix should be wholly positive for mouse markers. The first three panels in Figure 4.5 show these samples stained with the mouse depletion cocktail and the streptavidin conjugated secondary antibody. The fourth panel shows a control experiment stained with only the steptavitin secondary. As can be seen here, the mouse depletion cocktail is efficient in some cases, identifying over 90% of the marrow sample, but detects less than a third of the cells in the heart sample, and is even worse with the flank tissue. This suggests that while this method is efficient at removing lymphocytes and blood cell contamination, it is not effective for removing the fibroblasts which make up the subcutaneous membranes around the tumour implantation site. In these PDX models most tumours show relatively limited vascularisation (and residual red blood cells helpfully appear to suffer severe attrition during the fixation process), however the tumours are embedded within layers of subcutaneous membrane. As such, the least efficient of these examples is likely to be the most representative of the real population of contaminating host cells. Based on this, and the highly variable reliability even in tissues where detection was typically high, this methodology was not taken forward.

This presents a problem since any mouse cells that are not detected by the mouse lineage antibodies will appear in the same space as the EpCAM$^-$ subpopulation, contaminating what is arguably the most interesting (chemoresistent and tumour initiating) subpopulation. This is a particular issue given that this population is expected to be small, and so a relatively small amount of murine contamination would significantly impact experimental data for this population.

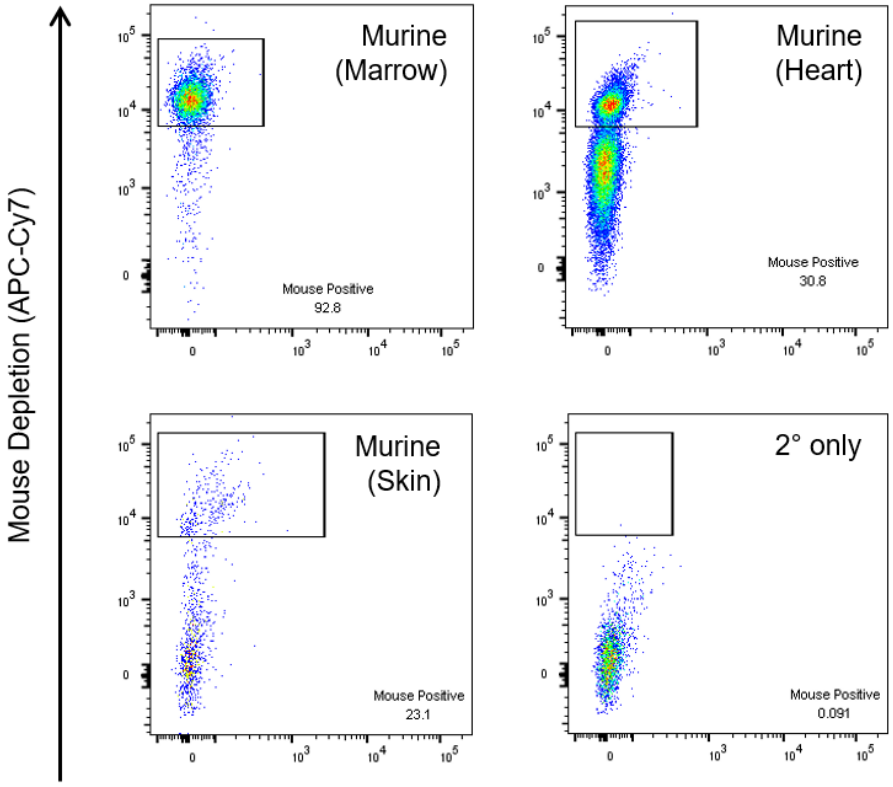


Figure 4.5: Depletion of mouse cells by mouse-specific antibodies varies drastically in its reliability. It is partiularly poor with the dermal tissue.

An alternative approach would be to use positive selection markers to isolate human cells. By using a ubiquitous human marker to isolate all human cells from debris and mouse cells, further work can be done on this purified population. To test this approach, methodology was adapted from previous work where this approach has been successfully used in similar conditions. Zena Werb's laboratory used CD298 to purify human cells from metastatic breast cancer cell-lines which had been used to create PDX models of metastatic disease, and the authors show mutually exclusive detection using CD298 and a mouse MHC marker (Lawson et al., 2015). Mark Shackleton's group report using an antibody to HLAs (Human Leukocyte Antigens) A, B, and C to purify human cells in PDX models of melanoma (Boyle et al., 2016).

This approach proved less successful with HGSOC xenografts as shown in Figure 4.6. As shown in the top right plot, these two markers struggle to clearly resolve mouse and human cells. The HLA antibody shows minimal almost no separation of the two cell types, and even where a handful of human cells do appear to have elevated signal, it is not clearly differentiated from the scattering of noise seen in the mouse-only sample. CD298 appears more effective, and while the two populations are not well resolved, there is clearly some separation occurring. Concerningly however, when a *very* conservatively selected CD298$^-$ population is assessed for human EpCAM (bottom panels) many of the cells from the xenograft are positive, demonstrating that even among the cells that are unequivocally CD298$^-$, many actually appear to be human. The absence of any such signal in the mouse-only control demonstrates that this is not a case of non-specific detection of mouse cells by the EpCAM antibody, but that CD298 must be entirely failing to detect a subset of human cells. Given that this project is interested in studying tumour cells with highly heterogeneous phenotypes, it is very unlikely that the subset of cells being missed by the CD298 gating strategy is a random selection, and thus it is very likely that important human subpopulations would be lost from the assay using this method. As such, using this strategy would likely eliminate a specific swathe of tumour cells, and as such represents a solution that is likely to be far more detrimental than the issue it is designed to solve.

While in the process of troubleshooting these methods, in an effort to develop a reliable separation methodology, a discovery arose somewhat serendipitously as a result of the use of ovarian cell-lines to generate xenografts. Cell-lines *in vitro* demonstrate a minimal level of phenotypic complexity, typically being wholly EpCAM$^+$. Occasionally a small EpCAM$^-$ subpopulation of $\sim$1% is seen, but it is hard to be confident it is not simply noise. As such, they are unsuitable for the type of studies in this project. However, during the initial stages of the project, when xenograft establishment was proving challenging, cell-lines were xenografted as a positive control to ensure there were no fundamental flaws in the engraftment methodology (as these lines are known to grow in immunodeficient mice). These samples grew successfully, albeit extremely slowly, and the subsequent tumours were used to optimise parts of the dissociation and flow cytometry methods in order to avoid wasting valuable PDX material. Surprisingly, when these tumours were dissociated, they showed significant phenotypic heterogeneity, similar to the PDX samples (see Figure 4.7). A similar

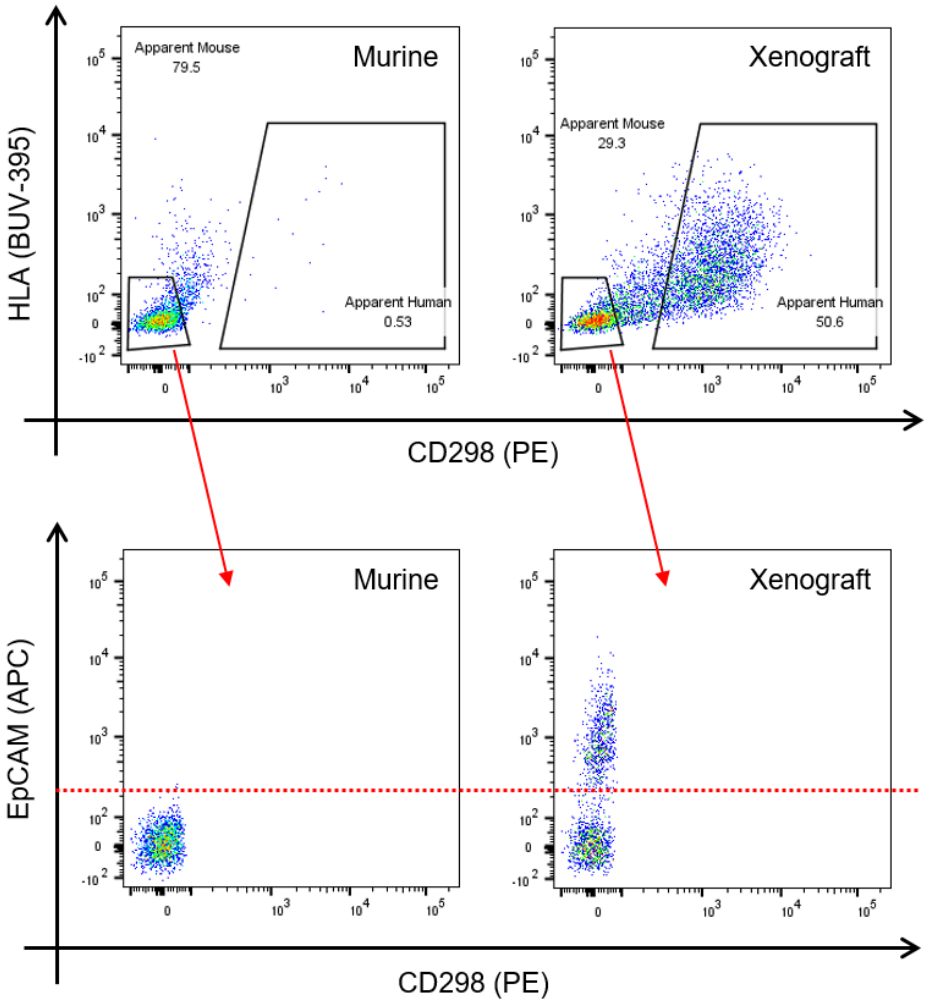


Figure 4.6: Pure murine tissue (left) compared against human HGSOC xenograft (majority human, some murine contamination; right). Both samples shown are incubated with antibodies against human EpCAM, CD298, and HLA. Bottom panels show EpCAM status of the $HLA^-/CD298^-$ gates from the respective figures above. The data show poor separation of the human material from mouse contamination (top right vs. top left). Far more concerningly, a significant proportion of the $HLA^-/CD298^-$ gate show positivity for human EpCAM that is not seen with pure mouse tissue (bottom right vs. bottom left), suggesting that these selection markers are failing to pick up a proportion of human cells.

spread of subpopulations was seen within these cell-line xenografts, however within each subpopulation there was much more uniformity between cells, making resolution of distinct subpopulations much easier. Alongside this work, efforts were made to troubleshoot the issues with mouse/human selection by titration of the antibody concentrations being used. While none of the alterations made to the selection antibodies proved useful, titrations of EpCAM and CD90 antibodies resulted in another interesting observation.

Up until this point, a binary $EpCAM^+$ / $EpCAM^-$ division had been assumed, based on Siru's previous studies: a positive fraction containing $EpCAM^+$ cells, and a negative fraction containing $EpCAM^-$ tumour cells and contaminating host cells. When EpCAM antibody was titrated, reducing the dilution factor from 1:12 to 1:50, 3 distinct populations could be resolved based on EpCAM profile. Within a PDX, this distinction may not have been immediately obvious due to the degree of noise arising from natural variation. However, the uniformity of the cells within each phenotypic subpopulation in the cell-line xenografts meant that the very clear resolution into three fractions based on EpCAM profile was undeniable. When co-stained for CD90, a fourth was also resolved (see Figure 4.7). These findings were then taken back to the PDXs where the same subpopulations were seen, albeit less clearly due to the higher level of noise.

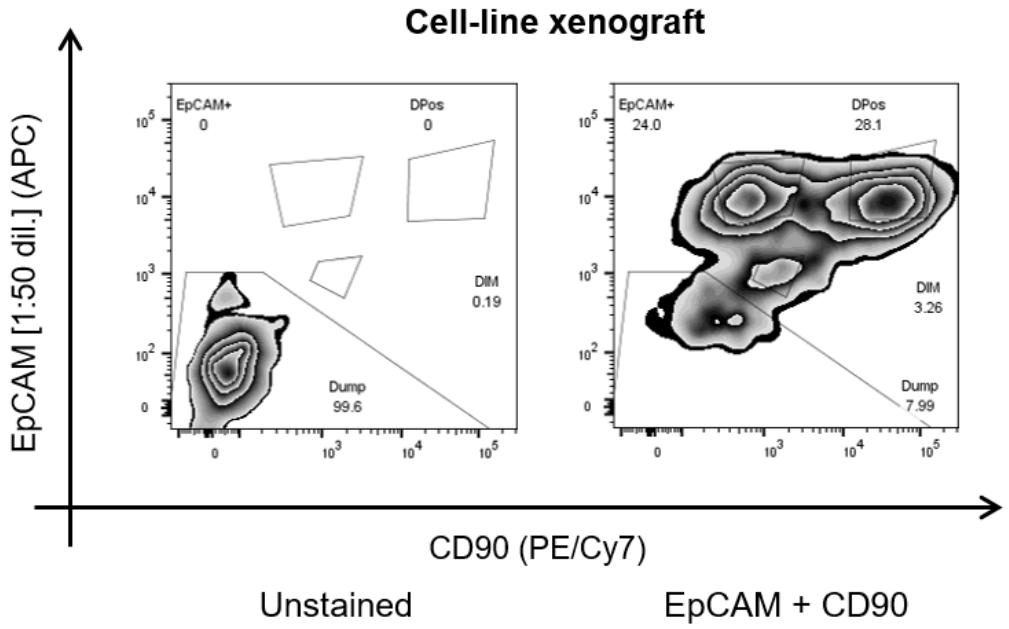


Figure 4.7: Application of a reduced concentration of EpCAM antibody resulted in the detection of multiple phenotypic subpopulations in cell-line xenografts.

When murine material, processed as previously, was treated with these antibodies, there was no elevation above untreated control, as per Figure 4.8. Xenograft material meanwhile shows

three distinct subpopulations (with a fourth subpopulation visible in some cases). The first is an EpCAM$^-$/CD90$^-$ subpopulation which appears at the same coordinates as pure murine material (and presumably represents murine cells and debris). The second is a strongly EpCAM positive subpopulation that is also CD90$^+$. The third subpopulation is found between these two, and appears weakly positive for both markers (subsequently referred to as EpCAM^{DIM}). In patient samples there is often a dense 'tail' extending diagonally from the double-negative population which looks very similar (see middle-left panel Figure 4.4); this 'population' is composed of debris of varying degrees of autofluorescence and is of no interest. Critically, the EpCAM^{DIM} subpopulation is distinct in that it shows fluorescence only in the EpCAM and CD90 channels, and typically resolves into a distinct area of cell density. Finally, there is sometimes a distinct fourth subpopulation, which retains the strong EpCAM expression of the EpCAM$^+$/CD90$^+$ (Double Positive) subpopulation, but is distinguished by low/absent CD90. This population is typically quite hard to resolve as a unique population, and may not have been indentified as a distinct subpopulation were it not seen so clearly as such in the cell-line xenografts. Initially it was assumed to be specific to particular samples, but has subsequently been observed in later passages of samples in which it was not previously observed. It is unclear if this subpopulation is actually absent in some cases, or if is present in all, but is too rare (or too weakly differentiated from the Double Positive) to be detectable as a distinct subpopulation in some cases. Interestingly it appears almost without exception in cell-line xenografts, although it is not clear if this is an indication that it is an artefact arising from long-term growth outside of the patient, or if it is simply that the uniformity of cell-line samples allows it to be reliably resolved from the Double Positive population.

4.3.3 Validation of subpopulation status by immunofluorescence

Detection above the threshold seen in stained mouse-only control tissue, provides evidence that the EpCAM^{DIM} subpopulation is human tumour and not contaminating mouse tissue. However, given the importance of this assumption to all future work, independent verification of this conclusion would be prudent. In order to do this, a xenograft tumour was digested (as per subsection 2.1.2), the subpopulations described above were sorted by flow cytometry (as per section 2.3), and the sorted single cells cytospun onto slides. Each population was then methanol fixed and stained for human p53 (as per subsection 2.4.2). A number of iterations were required in order to adapt existing methods for this experiment. Initial attempts to use pre-fixed cells were unsuccessful because the cells would not adhere well to the slide during the cytospin. After cytospinning of unfixed cells proved successful, 4% paraformaldehyde (PFA) and 10% neutral buffered formalin (NBF) were tested as possible suitable fixation reagents prior to antigen retrieval. Both broadly worked, but either appeared to inhibit subsequent IF detection, or resulted in cells losing adhesion during subsequent antigen retrieval. Subsequently fixation with methanol was tested, and while this still caused loss of cells using normal work-up, control slides (fixed and stained without retrieval to demonstrate successful adhesion) showed strong p53 staining, suggesting methanol

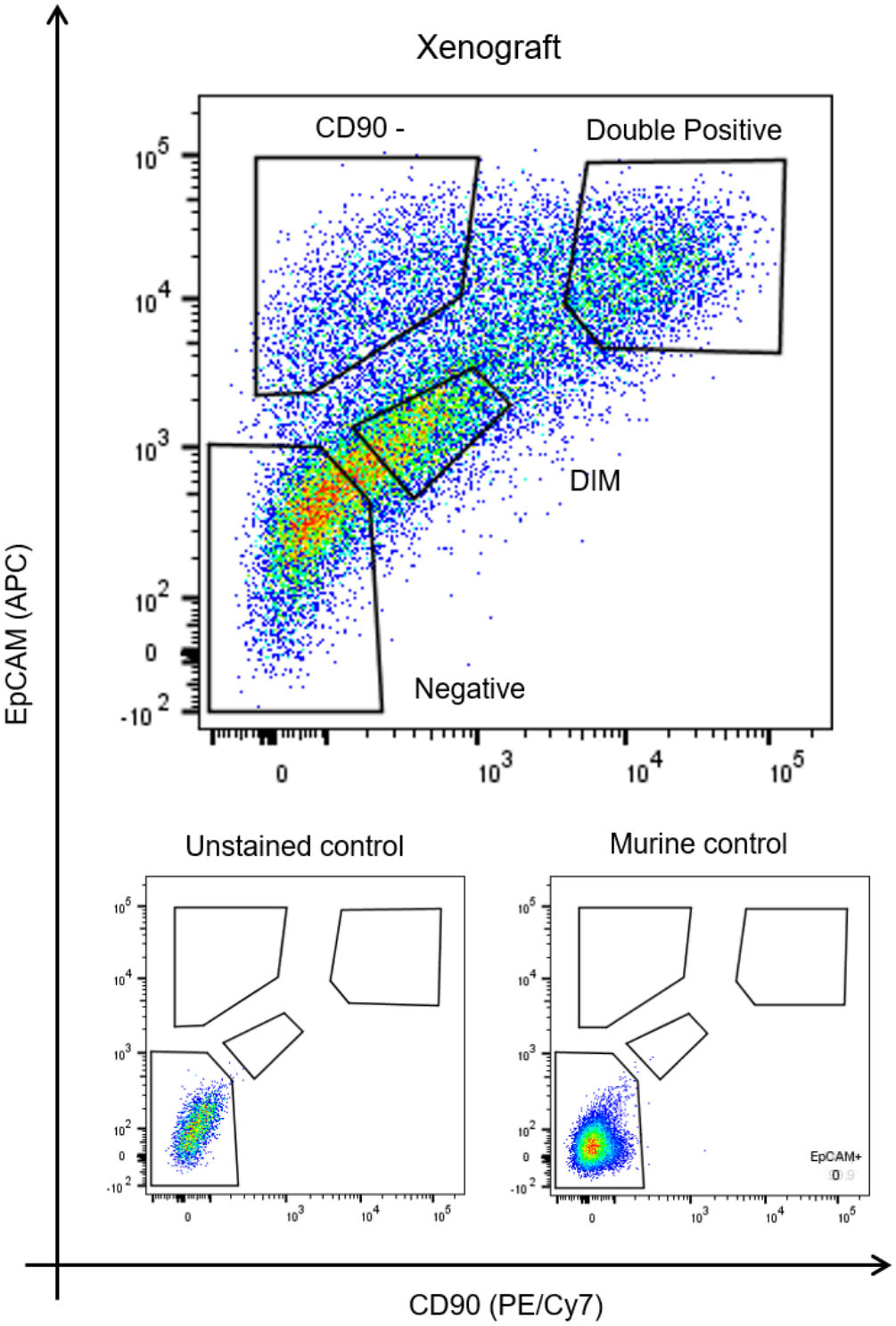


Figure 4.8: Flow sorting of whole xenograft by EpCAM and CD90, and division by subpopulation. A xenograft sample not stained with antibodies, and a stained sample of mouse material are shown as controls.

treatment is sufficient to permeablise cells and allow allowing p53 to be detected without an antigen-retrieval step.

The images shown in Figure 4.8 show cells from an OV04-449 xenograft imaged by confocal microscopy on a Leica SP5. As seen here, both the Double Positive subpopulation, and the $EpCAM^{DIM}$ subpopulation stain strongly for p53 (indicating they are human and HGSOC), while the $EpCAM^{-}$ subpopulation shows no signal (above background autofluorescence) as would be expected for mouse cells not expressing p53. Unfortunately, this method appears only to be able to detect the full-length protein, and the majority of the established xenografts for this project have mutations that induce premature termination of protein synthesis. These samples show no signal in any compartment, limiting the use of this test to the single sample shown (which was the only xenograft model with full-length p53 which had been developed at the time these experiments were performed). This lack of detection is likely to do with the necessary use of methanol as a suboptimal fixative. In particular, we can be confident that it is not related to the specificity of the antibody, as the same clone has been used successfully in subsequent IF experiments, where it is sensitive to p53, regardless of protein truncation. However, the OV04-449 result was replicated in a xenograft of the cell-line CIOV1 (derived from another HGSOC sample in the Brenton Laboratory by Maria Vias; manuscript in preparation), which shows the same set of phenotypic subpopulations, and a similar outcome from p53 staining. The cell-line result is valuable in that cell populations are well resolved within cell-line xenografts, and this provides additional confidence that the observed results are coming from the expected subpopulations, and not as a result of any 'bleeding' from one subpopulation into an adjacent one.

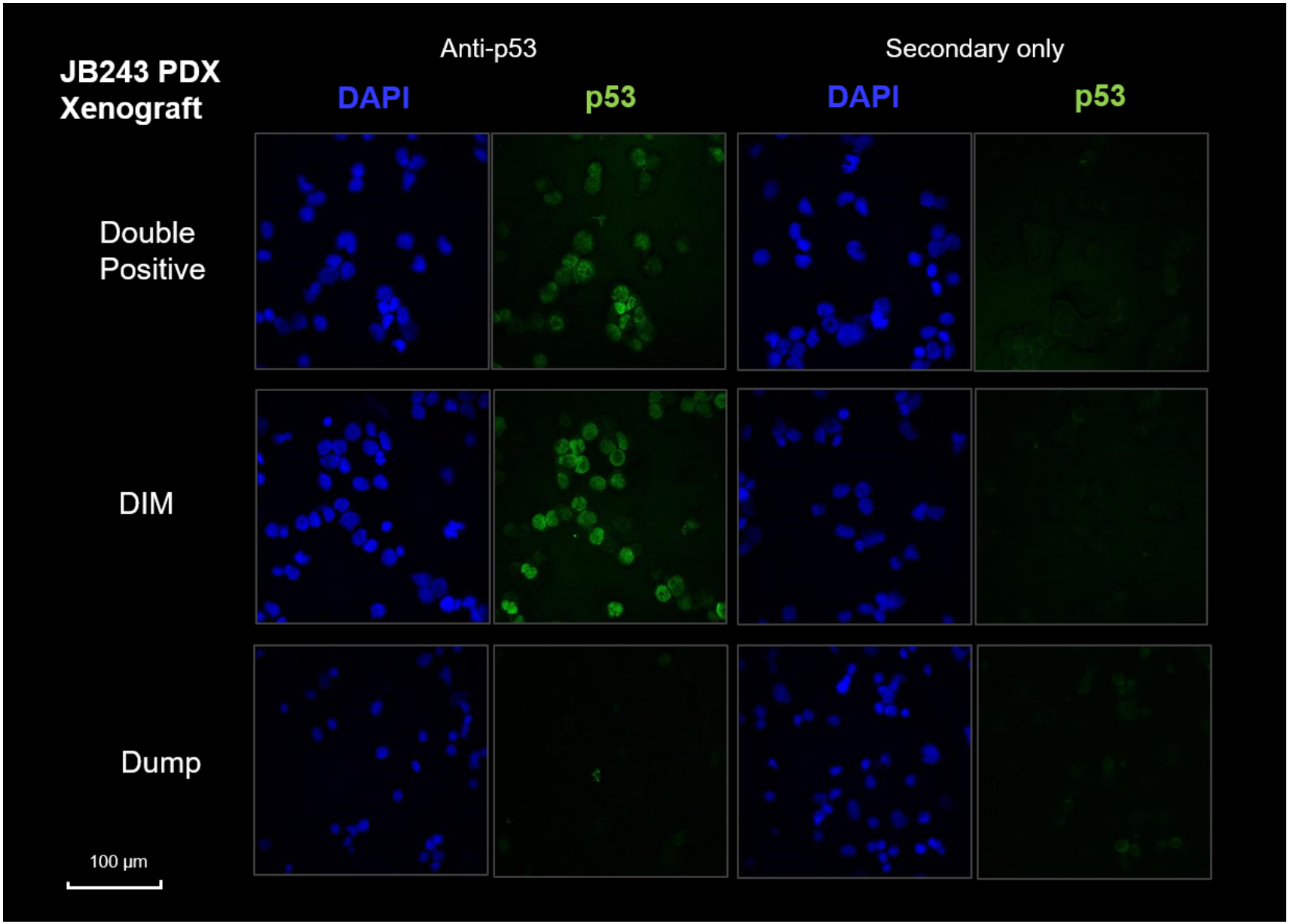


Figure 4.9: Confocal imaging of cytpospun cells from the three subpopulations, along with equivalent secondary-only controls.

4.4 Conclusions

Further work using the established xenograft models is dependent on the ability to reliably get from a xenograft tumour to distinct phenotypic subpopulations, via efficient dissociation, elimination of contaminating host cells, and analysis and resolution of phenotypically distinct subpopulations within the tumour as a whole. The work outlined above provides an optimised protocol for dissociation of tumours, a recapitulation of Siru's markers suitable for the xenograft context, an assessment of various methods for the depletion of mouse material, the final approach to depletion of mouse cells, and validation of the reliability of this method. The consistent set of subpopulations observed in the xenograft context is also outlined.

Optimisation of the dissociation protocol allows for higher yields of cells, which is particularly valuable when tumours are small. In addition, the process of maximising cell yield also minimises unnecessary cell death due to excessively harsh protocols. Such attrition may be biased against less resilient cells, and so optimisation here also improves how representative the resulting cell solution is of the original tumour. While it is unfortunate that podoplanin cannot be practically used in the xenograft context, CD90 provides a demonstrably reliable alternative. This marker was part of Siru's original screen, and was eliminated solely because it was redundant when podoplanin was also being used. The data shown here demonstrate that the $PDPN^+$ fraction of the patient sample falls exclusively within the $CD90^+$ fraction, suggesting that CD90 is present wherever podoplanin is found (although its use may also result in the capture of a fraction of $PDPN^-$ cells too).

The lack of suitable independent mouse/human discriminant markers is not optimal, however all of the methods tested showed serious issues in this context which make them unsuitable for use. Ideally a ubiquitous marker like p53 would be used to unambiguously separate tumour cells from normal mouse (at least in the majority of HGSOC samples where p53 is found to aggregate). While attempts were made to detect p53 by cytometry, the method was ineffective in isolating tumour cells due to negligible signal obtained. The method ultimately used, while far from optimal, is nonetheless adequate to the task, and demonstrates efficacy under controlled conditions. The phenotypic subpopulations, detected by human-specific antibodies, can be presumed to contain human cells, as similarly stained samples of representative mixed murine tissue do not extend into this region of the graph. In addition, isolation and staining of these cells for p53 by immunofluorescence reveals that these cells stain strongly for human p53, indicating that the selection process has been effective and these are indeed HGSOC tumour.

With an optimised protocol in place for going from patient tumour, to passaged xenograft, to phenotypic subpopulations separated by flow analysis, these subpopulations can now be experimentally interrogated in earnest.

Chapter 5

Characterising proliferation within phenotypic subpopulations

5.1 Introduction

With a set of established xenograft models available and an established process for isolating phenotypic subpopulations by flow cytometry, it becomes possible to probe the dynamics of tumour heterogeneity in intact tumours. An obvious first question is 'what are the *in situ* proliferation characteristics within each of the phenotypic subpopulations?' If the data from Siru's previous work is representative not just of tumour initiation, but also of growth in established tumours, then it would be expected that the $EpCAM^{DIM}$ subpopulation would be the only actively proliferating compartment, and the $EpCAM^{+}$ subpopulation would act simply as a bulk pool into which the $EpCAM^{DIM}$ subpopulation feeds new cells. On the other hand, if the Dual Compartment model is accurate, both subpopulations would be driving proliferation from within their own compartments. With a good *in situ* model, this question is now practical to empirically test. Xenograft-bearing mice can be dosed with nucleoside analogues, which will be taken up by any cell that divides during the dosing period. By following the rate of uptake into populations over time, the division patterns at the level of the population can be determined, and the division kinetics of each population characterised. Such experiments can offer significant understanding of the proliferation patterns within an *in situ* tumour. Notably what these experiments *cannot* do is determine division at the level of a single cell. The pattern of division can be used to support a Cancer Stem Cell or a Dual Compartment model, and potentially determine that significant proliferation occurs in the manner of one model or the other, but it cannot be used to completely exclude the other occurring at a minimal level.

Measuring uptake of nucleoide analogues in an *in situ* tumour, at the precision required here, is an invasive process that requires the killing of the animal and the removal and dissociation of tumours. Thus in order to conduct 'studies over time' (in reality a series of parallel experiments in which samples are harvested at different time-points), or experiments with different dosing conditions, the ability to produce cohorts of equivalent models is required. It is for this reason that the consistency of serially passaged cohorts is so valuable, and why in practice they represent a significant improvement over first generation xenografts.

Effectively identifying trends in proliferation necessitates ensuring that the experimental systems used are suitable and free from significant confounding variables. The suitability of the approaches used are validated in subsection 5.3.3 by assessing potential confounding factors, and demonstrating that these do not have a significant impact on the experimental results.

5.2 Methods

A summary of the primary method for data collection is shown in Figure 5.1, with detailed methodology in chapter 2.

By measuring the proportion of cells which have taken up the nucleoside analogue in replicate tumours, harvested at different time-points post-dosing, it is possible to infer the type of division based on dynamics of marker accumulation (which will only be incorporated to a significant degree via semi-conservative replication).

When plotting these data, rather than showing increase in cells positive for the analogue over time, the inverse is plotted (showing the drop in the percentage of cells remaining unlabelled over time from 100% at $t = 0$h). While perhaps counterintuitive, this makes calculating stochastic mechanics far easier since the resulting curves become simple exponential decay functions. Some hypothetical examples of possible patterns of results (and descriptions of the are proliferation mechanics that underlie them) are presented in Figure 5.2. The purple line shows a subpopulation that is not only quiescent, but which is also not being repopulated by another subpopulation (as either would lead to accumulation of labelled nucleosides); this population will simply not take up the analogue. The green line shows a subpopulation that is proliferating in a defined rather than a stochastic manner; in this case, all cells are dividing within a specific time-window, indicating that all cells are continuously dividing at a regular rate. The blue and red lines show stochastic division patterns; over a given time period, a certain fraction of all cells will divide; over the same period, that same proportion of the remainder will in turn divide, leading to an exponential decay in unlabelled cells. In this case the blue subpopulation is taking up label less quickly indicating a slower stochastic division rate. The orange subpopulation appears to have a mixed decay regime, indicating that it is likely made up of two distinct sub-subpopulations. While initially it appears stochastic, it tails off to an asymptote at around 60%, indicating that 60% of the

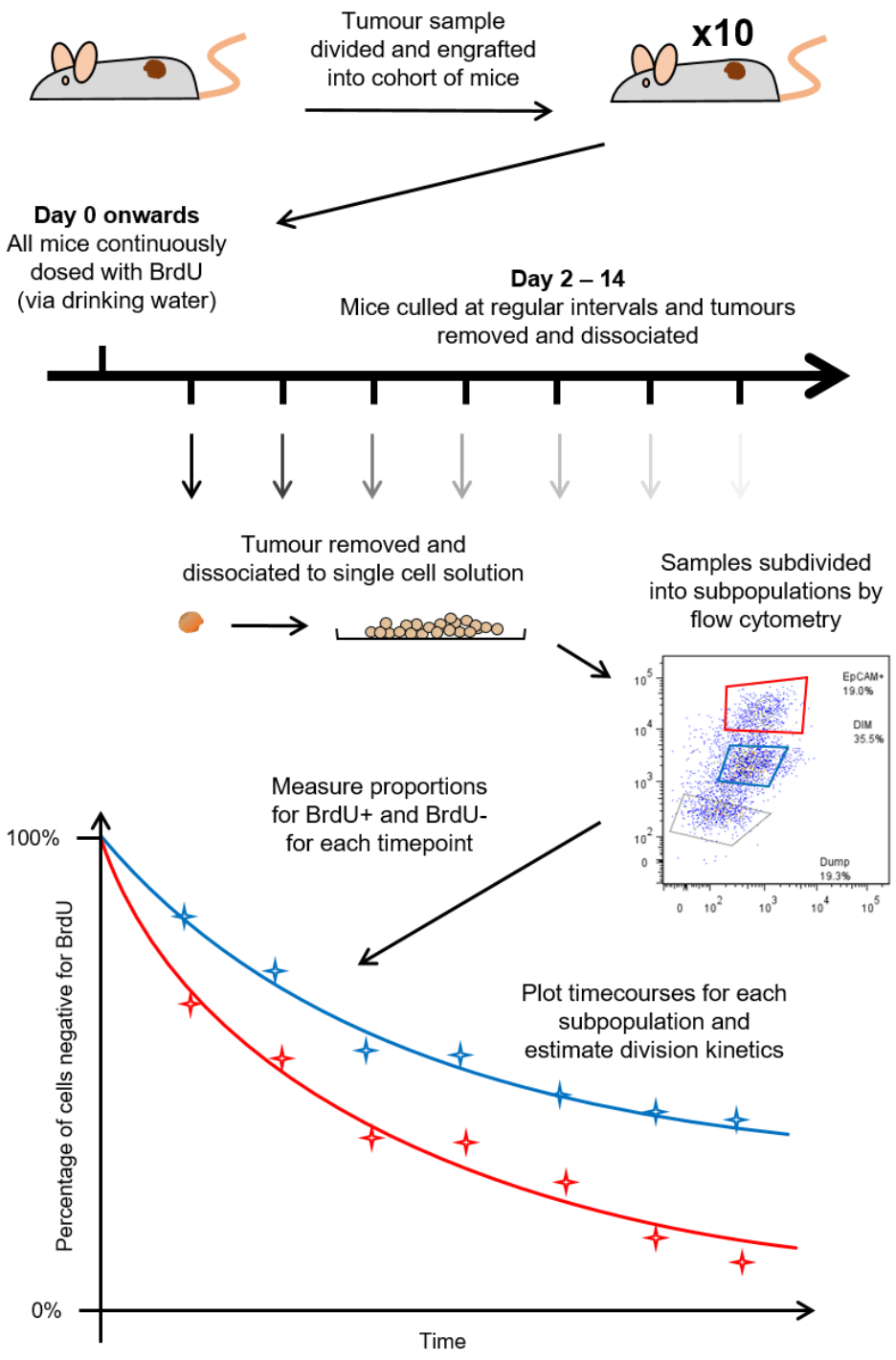


Figure 5.1: Method for estimating division rates and assessing differential proliferation rates within different phenotypic subpopulations.

subpopulation is quiescent. It is important to note that a fixed asymptote assumes that the subpopulation is not expanding significantly, while in practice this assumption will not be valid (and in reality this line would taper downwards as the quiescent subpopulation became a fixed pool in an increasingly large total subpopulation size).

Such plots not only show the dynamics of cell division within a subpopulation, but also allow division rates to be inferred. For cell populations with fixed division windows, the window is simply given by the intercept with the x-axis. For stochastic division, the plot should form an exponential decay curve, and the median division rate can be determined from the gradient of a logarithmically transformed plot.

A nucleoside analogue labelling time-course, run over the order of weeks, is efficient for determining the rate at which each of the subpopulations is being populated. What it is not easily able to do is determine which of the subpopulations is responsible for this proliferation. A rapidly cycling, self-supporting subpopulation is indistinguishable from a subpopulation populated by a progenitor of a different type, since both will see newly generated marked cells accumulating in this pool. As such, a second experiment is required to determine which subpopulations are responsible for driving division itself. This experiment is almost identical to the first; all that needs to be changed is the timescale on which it is conducted. Rather than taking up a tracing analogue over the course of days and weeks via drinking water, a couple of mice in a cohort can be retained, and treated with the analogue for only a couple of hours before they are killed and tumours extracted. As shown in Figure 5.3, this allows any cell that is in S-phase at the time of dose administration to incorporate the analogue. Since very few of these cells will have the time to go through mitosis, this method marks only the cells that have entered into cell cycle, rather than their daughter cells, some of which may actually be quiescent. Relative division rates between populations can be estimated from relative rates of marking, and with an estimate of S-phase length, absolute division rates can also be inferred. As the short-term window is so limited, mice are unlikely to drink from their water bottles in this time (even before considering possible delay in uptake to the bloodstream and from there to tumour), and so a more direct administration is required. In order to ensure the analogue reaches the tumour, and that there is sufficeint time for it to be taken up and incorporated into the DNA of cells, the analogue is instead injected intraperitoneally. This method has been shown to be successful in previous studies (Giraddi et al., 2015).

Comparison of the estimates derived from the long-term and short-term nucleoside analogue incorporation experiments allow subpopulation dynamics to be inferred. If all populations are self-supporting (the working hypothesis in this case) then it would be expected that rate of entry into cell cycle (measured by the short-term experiment) and rate of new marked cells entering the subpopulation (measured by the long-term experiment) should be the same. If the Cancer Stem Cell hypothesis is correct, the $EpCAM^{DIM}$ subpopulation would be strongly marked in the short-term experiment, while the $EpCAM^{+}$ subpopulation would be essentially unmarked as proliferation is restricted to the $EpCAM^{DIM}$ subpopulation. However,

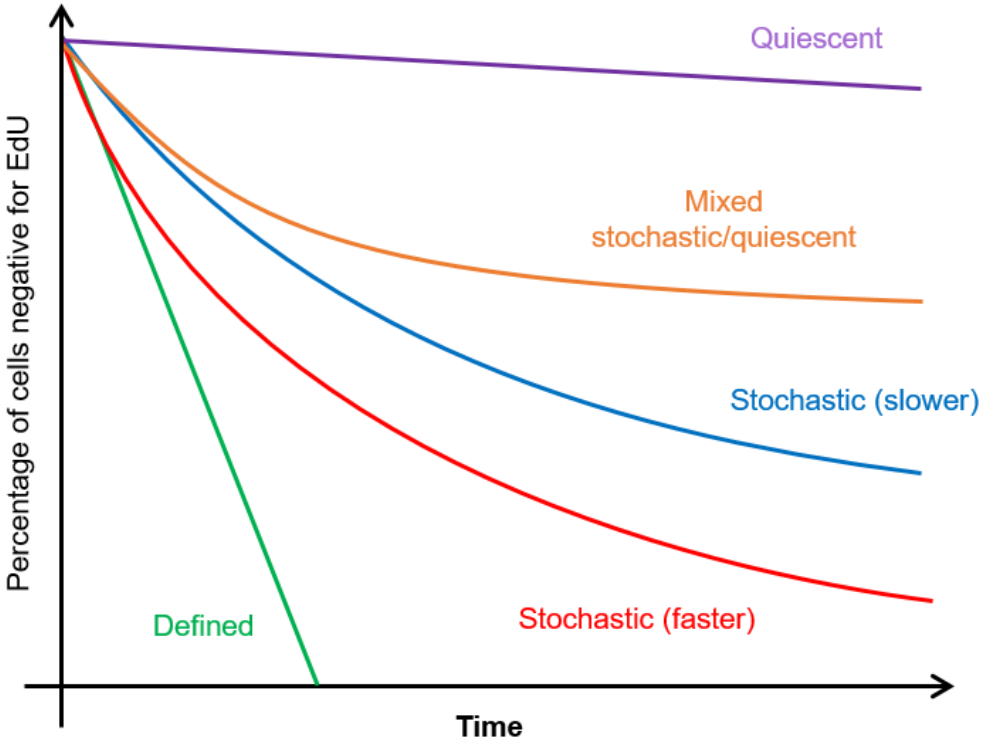


Figure 5.2: Examples of possible expected datasets, given hypothetical proliferation dynamics. These include quiescent (purple), defined cycling-time (green), stochastic (blue and red), and mix of quiescent and stochastic cells (orange).

the long-term experiment would show stronger marking of the EpCAM$^+$ subpopulation, since the majority of daughters will become quiescent EpCAM$^+$ cells.

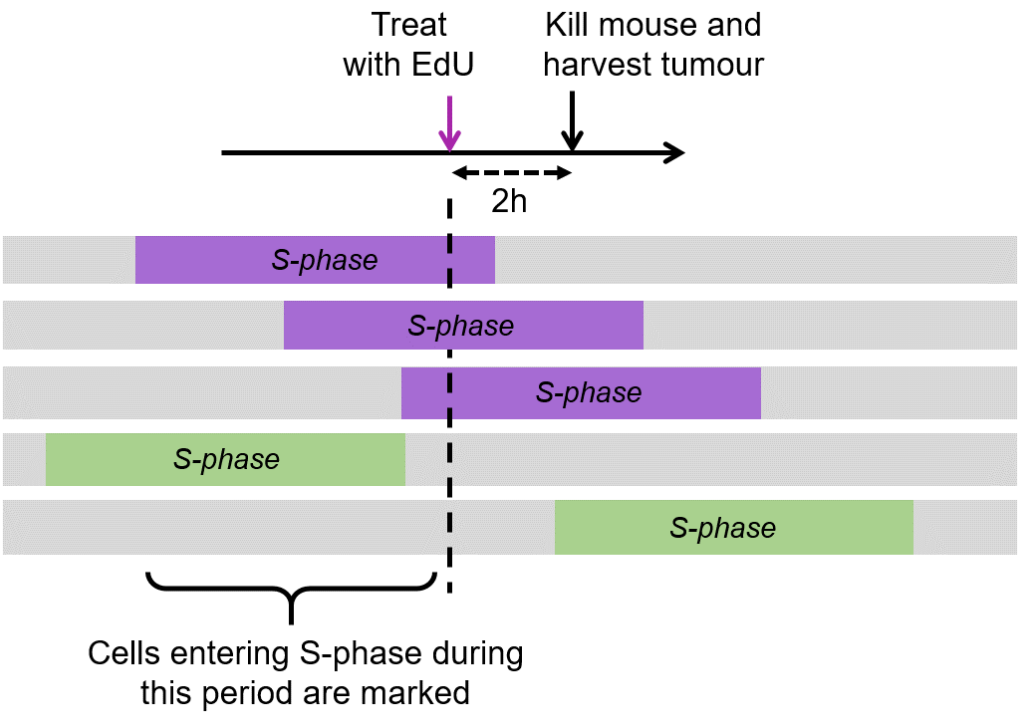


Figure 5.3: Explanatory summary of the short-term nucleoside tracing experiment. Tumour-bearing mice are dosed 2 h before the mice are killed and tumours harvested. Of the five cells shown, three are in S-phase during the dosing, taking the analogue up and incorporating it into new DNA. These cells do not have time to divide before harvest and measurement, so only the parental cell will be marked. The last two cells respectively finish S-phase before dosing, or begin too late to take up a meaningful degree of the analogue prior to tumour harvest. This leaves a window approximately the length of S-phase; thus any cell dividing within this time period will be marked with the analogue, revealing proliferation rates within each subpopulation.

5.3 Results

5.3.1 Iteration of experimental design

The initial experimental plan was to treat cells with the nucleoside analogue 5-bromo-2′-deoxyuridine (BrdU), physically sort cells by subpopulation using flow cytometry, then cytospin these cells onto microscope slides and determine the proportion of BrdU$^+$ cells using immunofluorescence. Pilot trials of this workflow proved challenging, because while BrdU

detection by immunofluorescence in fixed tissue was trivial, established antigen retrieval methods (in this case opening the DNA enough to expose the BrdU to antibody binding) proved sufficeintly harsh to cause cytospun cells to detach from the slide. It proved difficult to optimise a method that both retained adhesion but also exposed the antigen (with the more promising fixation/retrieval conditions introducing additional problems with non-specific detection).

An alternative method is to use 5-Ethynyl-2′-deoxyuridine (EdU). Unlike BrdU (and other similar halogenated analogues like CldU and IdU), EdU is modified not with a halide but with an ethynyl reactive group. Rather than being detected by an antibody specific to the structure (including the specific halide), EdU can be detected by addition of a detection tag of choice bound to a reactive azide group which reacts with the ethynyl. Since the probe in this case is a small molecule rather than an antibody, cells do not need to undergo the aggressive treatment required when using BrdU, as the reagent is smaller and can access the DNA far more easily. Since the process is chemical rather than antibody-based, and azide and ethynyl groups are not naturally occurring, there is little or no non-specific detection, something that had proved a significant problem with pilot BrdU work. Perhaps most significantly, the chemistry involved is compatible with flow cytometry which means that EdU detection can be incorporated into the flow cytometry process. This allows subsequent steps to be eliminated, removing a number of complex additional stages with potential to introduce variability. Cells now need only to be analysed using flow cytometry rather than flow sorted, and there is no need for cytospin, antigen retrieval, immunofluorescence, or imaging. Many of the problems with the BrdU-based method arose primarily out of the additional steps – BrdU detection was challenging mainly because of the need to cytospin samples, not because the assay itself is difficult. As such, the EdU method's elimination of these is a huge benefit. It also means the assay can be performed in a day rather than two, which makes lengthy time-course experiments much more logistically practical. Finally it eliminates the need for imaging, analysis and quantification of cytospins, which entails complexity, time outlay, and a major source of possible variability in results.

Detection of EdU was performed using an off-the-shelf kit from Invitrogen (Click-iT Plus EdU Alexa Fluor 488 Imaging Kit) which provides a conjugated fluorophore/azide detection reagent, and optimised fixation and permeablisation buffers. Minor adaptations were made to the provided protocol (see chapter 2) due to experimental design and the findings of pilot experiments.

The given protocol does not appear to account for the need to split samples to provide controls for cytometry, and this had to be rectified. The provided protocol calls for antibodies to extracellular antigens to be added prior to fixation, presumably due to the risk of degradation/deformation of antigens during fixation and permeablisation. Pilot work showed this method appears to work poorly in this case, with weak signal from extracellular antigens (presumably because antibodies are washed off in the multitude of subsequent steps, or their binding weakened by the reagents being used). Addition of antibodies after EdU detection

appears to resolve this problem with no adverse effects. The signal from extracellular antigens is similar to unfixed cells, and those fixed but not treated with the EdU buffer and reagents. The only meaningful effect is that fixation itself appears to cause a small shift in fluorescence for these markers, but this occurs to all cells regardless of fluorescence status, translating the entire graph a fraction away from the origin and introducing a small amount of noise. This is not a problem, but it is important to be aware of the effect when comparing fixed and unfixed samples as population gates must account for this small shift. Finally, the recommended cell numbers for this experiment appear to be far too high – using 10 million cells is recommended, but empirical experimentation in this context suggests even 1 million (per 'test' of the EdU kit) is too high to allow for reliable detection over a full time-course experiment (where labelling eventually approaches 100%). After two weeks of continuous dosing, it appears sufficient EdU is incorporated that the detection reagent becomes limiting, as the previously largely binary division degenerates into a smear where the true EdU^+ population is hard to accurately determine. Optimisation suggests reducing the cell number to 0.5-0.25 million cells is more appropriate for the purposes of the time-course experiments being conducted here. At this level, detection is reliable until at ~90% of cells are marked. Beyond this, reliability degrades, and it is hard to get accurate values despite increased detection reagent (both by concentration and volume). It is possible this effect arises from another cause – levels of EdU detection >90% are only seen in the last few time-points, and it may be that after a fortnight of exposure to BrdU, even cells which are not proliferating may have slowly assimilated enough EdU to begin to look slightly positive, destroying the ability to reliably resolve the EdU positive and negative fractions clearly.

The gating strategy used in the experiment is shown in Figure 5.4 and Figure 5.5. Crude removal of debris and multiplet cells is by routine means as shown. Removal of dead cells was initially performed using a fixable live-dead violet dye from Invitrogen (since cells are to be fixed, the default approach of adding DAPI as a marker cannot be used since even the cells that were live will have been permeablised – this fixable marker is added prior to fixation, permeates only dead cells, and is not washed out by subsequent fixation). After an unrelated experiment revealed a large amount of debris which did not absorb the dye, this protocol was altered (for the last three time-courses) to also include PI (propidium iodide), a DNA intercalating agent, which allows nucleated cells to be differentiated from anuclear debris and erythrocytes which are not of interest for time-course studies.
The alternative methods are shown in the bottom row of Figure 5.4, and both figures presented are typical (though from distinct experiments). As seen in the latter, there are few anuclear cells (presumably many of these will be erythrocytes which will be discarded later anyway given their mouse origin), and nothing similar to the aberrant sample, rich in large anuclear debris, has subsequently been seen.

The addition of a step to remove possible such debris was the correct choice in the circumstances since it ensured proper rigour and control of an unknown variable which could potentially have compromised the validity of the data being collected. However, it had consequences in that the fluorescence of PI bled into the PE-Cy7 (CD90) detection channel, and

impacted detection quality for CD90, in particular making it more challenging to separate the $EpCAM^{DIM}$ subpopulation from the true non-human $EpCAM^{-}$ subpopulation. While compensation can be applied to estimate and remove this effect, compensation always adds additional uncertainty, which can make the difference between distinguishing two subtly different subpopulations and being unable to do so.

In hindsight, given that it is now clear that anuclear debris is not a significant concern, the later data is likely to be less reliable than that collected using the original method. Were similar work to be conducted again, I would not stain the final experiment sample with PI and instead add an extra sample to the flow process in order to stain for PI in isolation. This would confirm the absence of extensive anuclear debris. Given the extreme rarity of this problem I believe it would be better to risk having to discard the whole sample in one or two cases than introduce a strong fluorescence signal to every sample and compromise the ability to resolve populations.

An alternative approach which would sidestep most of the drawbacks of the method used, would be to swap the two the fluorophores being used. PI and DAPI are functionally equivalent DNA markers that effectively measure cell permeability, and Invitrogen offer the live-dead fixable dye with a range of wavelengths. By replacing PI and violet live-dead dye, with DAPI and a live-dead dye in the red channel, the setup remains broadly the same but the markers would read out in opposite detection channels. Since the live cells that are actually of interest to the experiment are nuclear marker positive, but fixable dye negative, they would now emit strongly in the violet channel, but minimally in the red channel. Even if the percentage bleed from the channel was as high as for P.I. (Invitrogen do not appear to publish spectra, but it would be surprising if a custom molecule leaked as strongly as PI in the far-red region) then the leakage would be a small percentage of 10^2 (fixable live-dead negative) rather than 10^4 (nuclear positive). This does risk introducing similar problems in distinguishing live and dead cells (fixable live-dead marker negative and positive) due to leakage from the CD90 channel into the live-dead channel, however this would be much less pronounced as the live-dead populations are well separated and CD90 is not as strongly expressed.

Once live cells were selected, they were then assessed (Figure 5.5) by their EpCAM and CD90 status as previously discussed, and the percentage of EdU positive cells calculated for each subpopulation.

Pilot tests were performed on cell-lines to validate that EdU and extracellular cytometry were working as expected, and that the methodology was adapted as previously noted. The system was then tested in xenografts to confirm suitability. With this done, two preliminary time-courses were performed on two groups of four mice, each group being implanted with PDX material from the same mouse (and each group containing one of two distinct biological samples). A third group was set up similarly with an HGSOC cell-line (PEO-4) rather than a PDX sample. A mouse from each group was killed on days 1, 3, 5, and 7 after the beginning of continuous dosing, and their tumours removed and assayed.

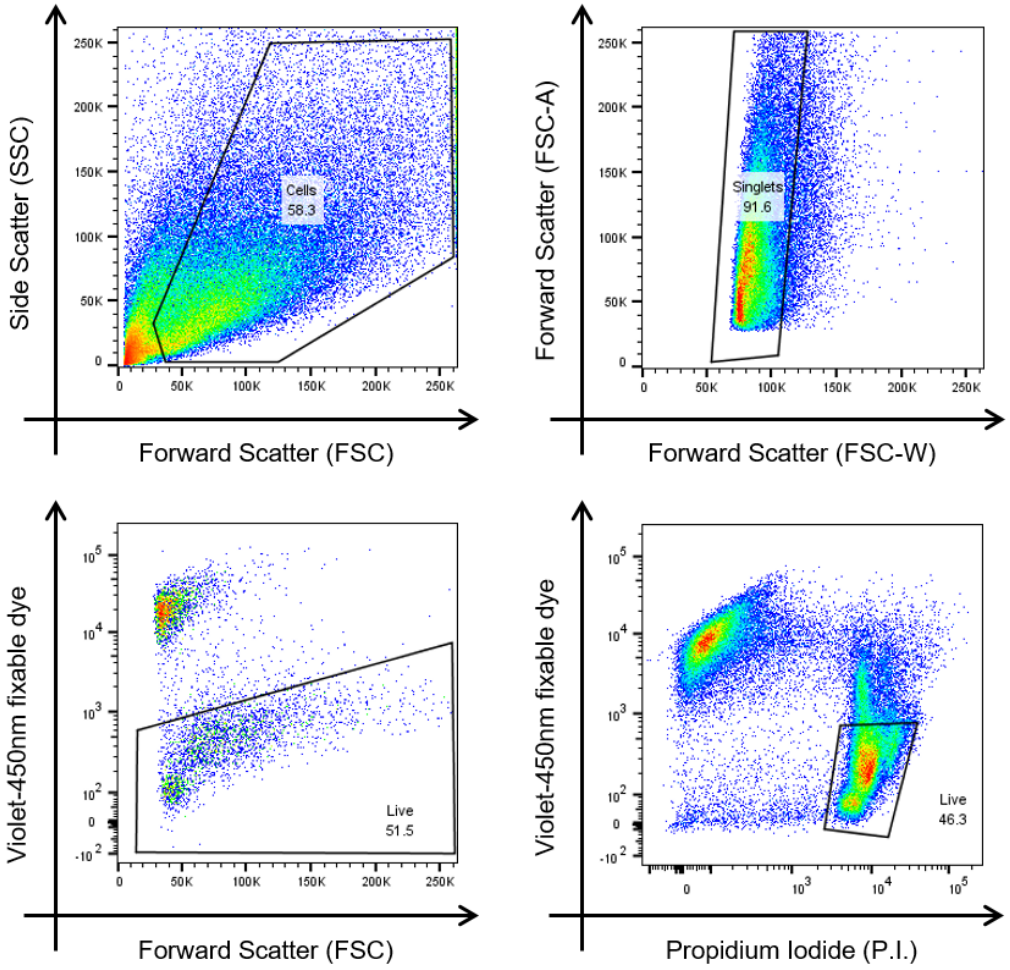


Figure 5.4: Analysis gates used to select out live cells in flow cytometry. Cells were selected from debris based on forward- and side-scatter properties (top left). Adhered pairs of cells within this cellular population were removed based on magnitude/duration of forward-scatter signal peak (top right; the discarded 'tail' to the right is multiplets of cells). Dead cells were removed using one of two strategies shown below. Initially dead cells were removed solely based on uptake of a fixation-resilient violet dye that only permeates dead cells (bottom left; dye is added prior to fixation). For later work a post-fixation propidium iodide nuclear stain was also added, and live cells deemed to be those not taking up the violet permeability dye, but staining with propidium iodide (bottom right). Propidim iodide is intercalated into DNA, meaning it only stains nucleated cells allowing anuclear debris and erythrocytes to be discarded.

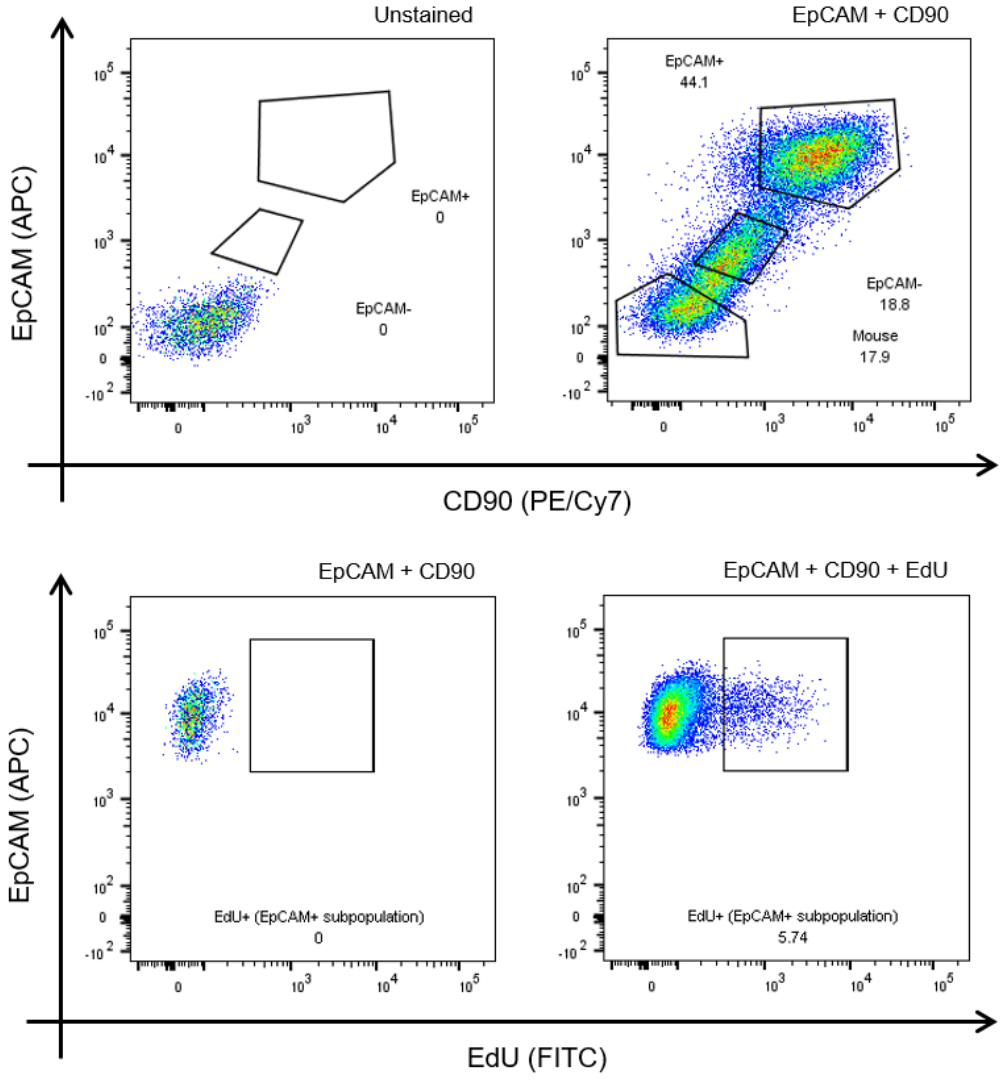


Figure 5.5: Gating used to select subpopulations (top; unstained control xenograft versus EpCAM and CD90 stained xenograft sample) and subsequent selection of EdU^+ cells from each subpopulation (bottom, EpCAM and EdU detection reagent vs. EpCAM alone).

The results were very encouraging, and the approach was expanded to include more biological replicates (xenograft tumours with different patients of origin), replicate time-courses (two cohorts of mice from the same sample of origin, but passaged from a distinct xenograft and not tested as part of the same batch, in order to investigate variability in results arising from the same original tumour). The numbers in each cohort were also increased to improve power. In addition, in subsequent time-course cohorts a few mice were set aside to provide matched short-term EdU dosing data for each time-course experiment. These mice were selected arbitrarily and the cohort size expanded commensurately. As such, the exact design of time-courses evolved gradually as more information was obtained, methods were improved, and the power of the experiment optimised. The biggest change occurred early, immediately after the two preliminary time-course experiments. Initially, the two tumours from each mouse were pooled, processed and assessed. However, since these tumours grow in relative isolation, they are functionally technical replicates, and pooling them only serves to reduce the power of the experiment. While it is reasonable to expect that various systemic effects (e.g. minor transient illness in the mouse, or the rapid growth of one tumour limiting the growth rate of another by competition) could have non-independent effects on both tumours, it should not ultimately affect the interplay between the subpopulations in any given tumour (given that signalling processes between tumour cells are likely to be extremely local in nature). While competition for resources might lead to one a tumour being smaller (this effect has been reported in some literature although no evidence for it was seen in this case) it should still retain the same proportions of subpopulations, and those subpopulations should still be dividing in the same way. Assessment of the resulting data shows that pairs of tumours from the same mouse show a similar degree of variation as the cohort as a whole.

With an increased cohort size being adopted in response to the success of preliminary time-courses, the question arose of the best placement of time-points in order to provide the most useful information. The general design used was a first time-point around 20 h (day 1), followed by subsequent time-points every 2 days. The 20 h time-point was adopted as a compromise between the value and risk of an early time-point. Given a stochastic division rate:

$$\text{rate of cell marking} = \text{fixed constant of division} \times \text{proportion of cells unmarked}$$

Therefore, the experiment is most sensitive early on when the fewest cells are marked. However, mice are not truly dosed at $t = 0$ h, but at the point they first drink from the EdU-laced water; this margin of error rapidly becomes insignificant, but at very early time-points would make data collection unreliable. Mice are nocturnal in nature, and the EdU-laced water was added at the end of the day, so mice are likely to begin drinking a few hours after dosed water is added. It can therefore be assumed that by 20 h the mice have drunk some water and had EdU entering their system for the majority of this period. However, the number of cells which have divided at this point is only around 10%, ensuring that an early high-power time-point is obtained. Subsequent time-points were every 48-96 h. While it would be possible to take time-points every 24 h, there is no particular advantage to

doing so since these values can be reasonably interpolated, and the logistics are challenging. Some time-courses were extended out as far as 18 days, but measurements at this extreme turned out to be of limited use due to two factors – firstly, the relative change in percentage of marked cells becomes very small at this point because the majority of cells are already marked; secondly, the technical limitations of the EdU detection method make it challenging to resolve accurate proportions of unmarked cells below a threshold of around 10%, making data collected after this time unreliable. As such, most experiments ran for around two weeks. Finally, cohorts were typically prepared with an extra mouse or two to account for the risk of attrition due to illness (NSG mice being at significant risk from fairly minor ailments), rather than risk rare attrition leading to missed observations. The vast majority of the cohorts survived without attrition, and the extra mice were subsequently used to add power to their experiment. In general, in these cases, two mice were taken for time-points around the 3-7 d mark, rather than one, in order to increase the power of these time-points. This is because experience showed that this was where the greatest variability in results was seen, and where additional replicates would be most useful.

In total 8 xenograft time-courses were performed on 6 distinct patient samples (with two samples used in two separate time-course replicates), as well as the one HGSOC cell-line.

5.3.2 Analysis of time-course data

The data collected from all time-courses were consistent with exponential decay, and an appropriate analysis process was applied. This is summarised in Figure 5.6, and a summary of all the long-term data for all nine time-courses is shown (in logarithmically transformed form) in Figure 5.7 along with the linear regressions used for analysis and subsequent calculations.

The $CD90^-$ subpopulation appears in only two PDX samples, and the one cell line sample.

Two-subpopulation analysis

For simplicity, the remaining six samples (in which there were only two subpopulations) were analysed first.

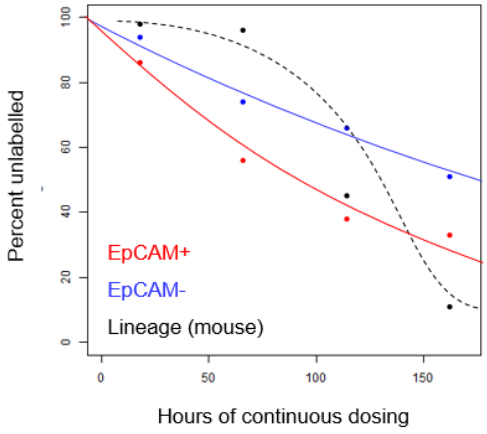
Long-term assays

Figure 5.8 shows the plots of these cohorts in more detail; the shaded uncertainties represent 95% confidence intervals, or two Standard Errors around each regression.

There are two key conclusions which can be drawn from these data:

Exponential decay plot

Initial plot of exponential decay curves showing loss of unmarked cells



Decay constant

The exponential decay equation (top) can be manipulated into the form of a linear graph ($y = mx$; bottom)

$$N = N_0 e^{-\lambda t}$$

$$\% = (N / N_0) = e^{-\lambda t}$$

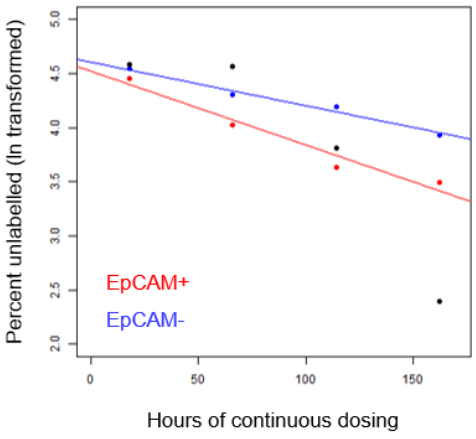
$$Ln(\%) = -\lambda t$$

Transformation

A graphic representation of this equation can be generated by logarithmic transformation of the data on the y-axis

Decay constant

The negative gradient of this linear equation gives the exponential decay constant λ which can be calculated by a linear regression



Average proliferation rate

The equation for calculating the half-life of an exponential decay (in this context the time required for half of cells to be marked, i.e. median division time) can be calculated from λ

$$\text{Median time to division} = ln2 / \lambda$$

Figure 5.6: Data collected in time-course experiments show exponential decay, allowing analysis to be based on the simple mathematics that underlie the process. The mouse lineage population (black) follows no meaningful pattern because it does not represent a closed stochastic system, and is presented here as contrast to the clean stochastic decay.

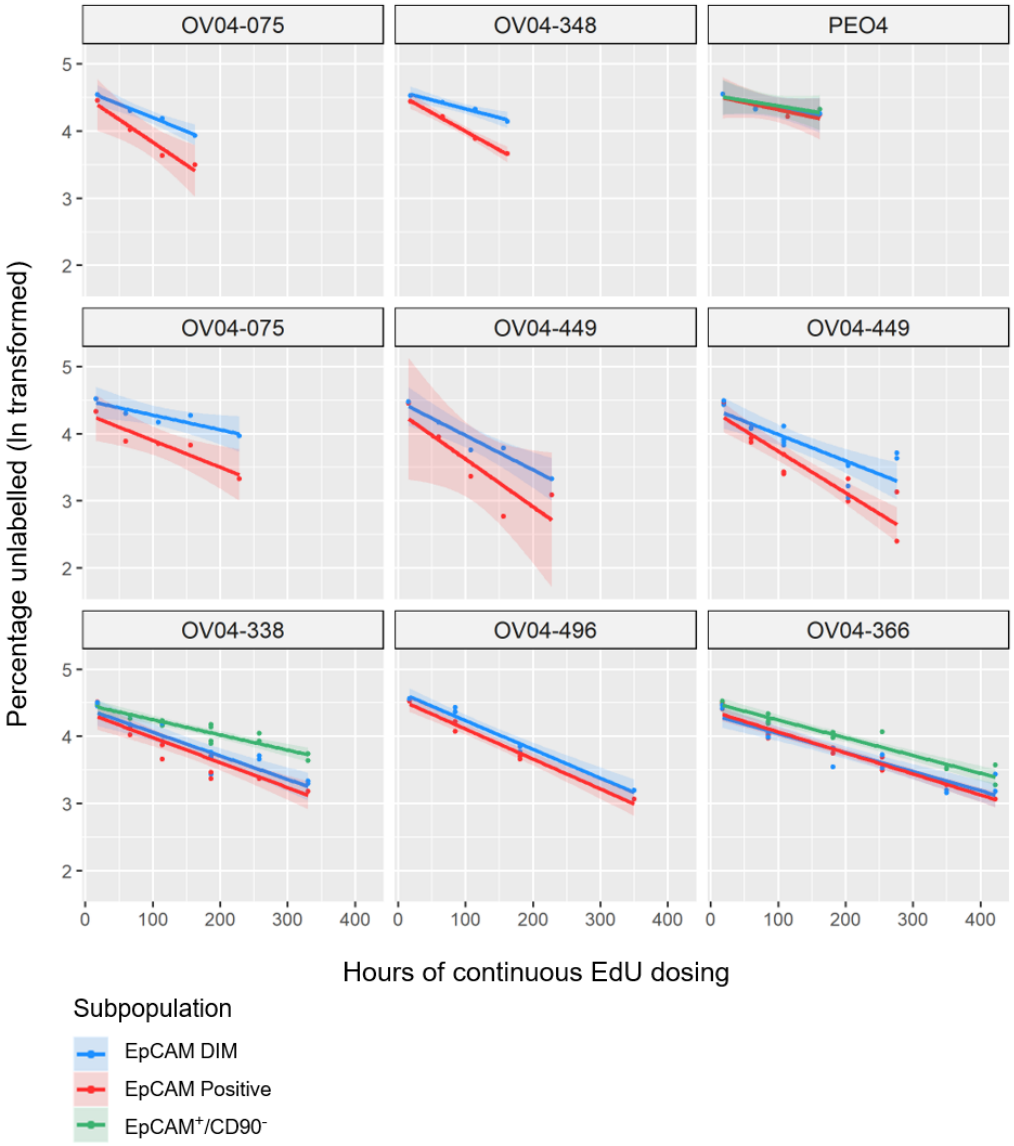


Figure 5.7: Plots of logarithmically transformed EdU uptake for all nine time-courses . Data points are shown for each subpopulation, along with linear regressions for each. The shaded areas represent 95% confidence intervals (calculated as mean ± 2 SE).

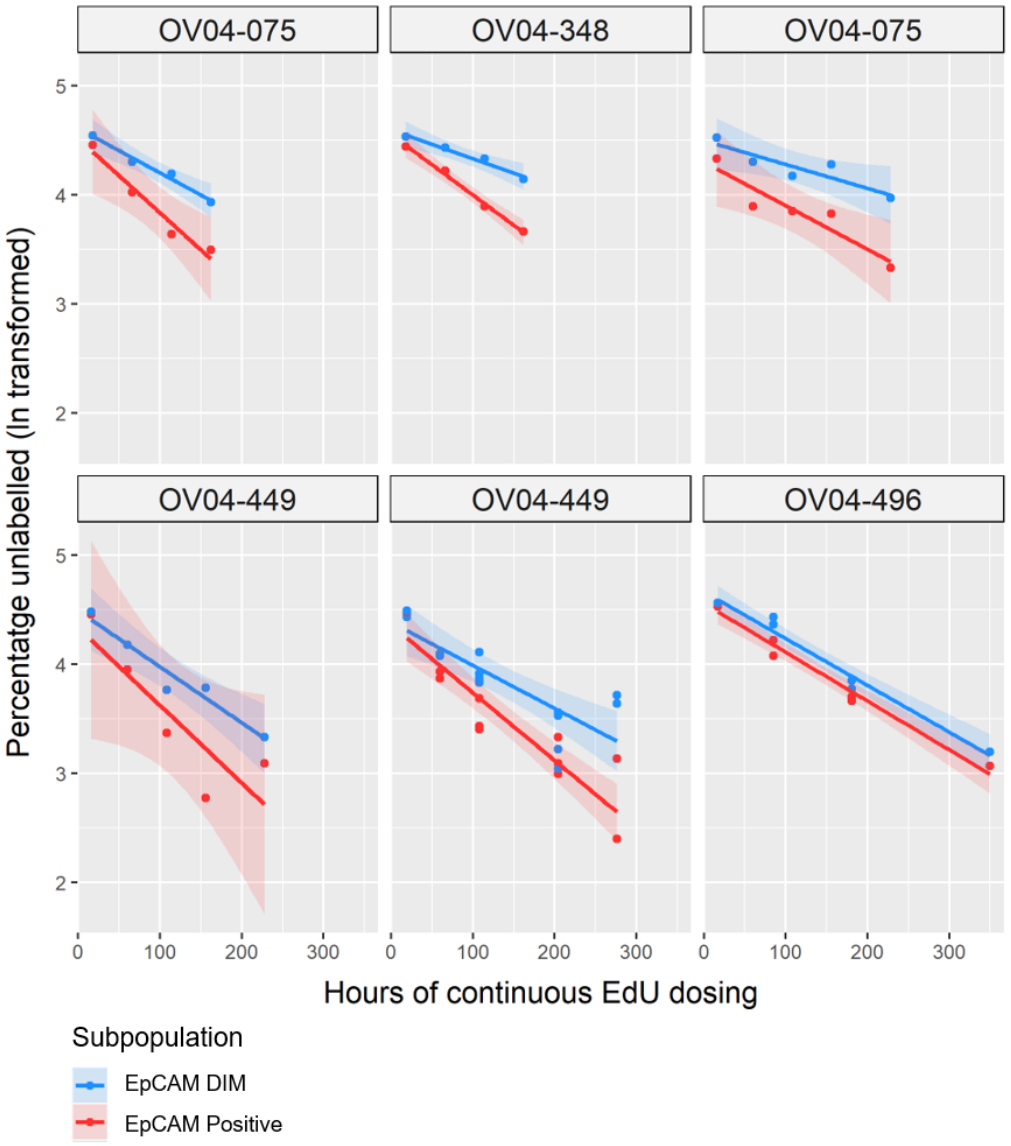


Figure 5.8: Logarithmic plots for samples with two subpopulations only. The shaded areas represent 95% confidence intervals (calculated as mean ± 2 SE).

- The data form exponential decay curves, which when logarithmically transformed become linear. This pattern is seen for both $EpCAM^{+}$ and $EpCAM^{DIM}$ populations. This indicates that both populations are proliferating in a stochastic fashion.
- The $EpCAM^{DIM}$ population is proliferating more slowly than the $EpCAM^{+}$ in all cases (though the relative magnitude of this varies).

These data allow the immediate exclusion of an $EpCAM^{DIM}$ cancer stem cell hypotheses that does not use $EpCAM^{+}$ transit-amplifying intermediaries, since the proliferation in the $EpCAM^{DIM}$ subpopulation (the proposed CSC subpopulation, and a much smaller subpopulation than the $EpCAM^{+}$) is not sufficient to account for the rate of EdU accumulation in the $EpCAM^{+}$ subpopulation.

A model where the entire $EpCAM^{DIM}$ subpopulation is proliferating (i.e. all $EpCAM^{DIM}$ are CSCs) would predict the $EpCAM^{DIM}$ subpopulation accumulating EdU quicker than is seen here in order to sustain the observed incorporation of EdU into the $EpCAM^{+}$ fraction. Alternatively, a model where only a small proportion of $EpCAM^{DIM}$ cells are CSCs and driving proliferation (while the rest are quiescent) is also excluded since it would require all EdU incorporation seen here in both subpopulations to be driven by a very small fraction of the $EpCAM^{DIM}$ population. This would require such a high rate of proliferation in this small fraction as to be biologically implausible given the rate of incorporation by $EpCAM^{+}$ cells.

The slower rate of division seen in the $EpCAM^{DIM}$ population, as compared to the $EpCAM^{+}$, suggests that these populations have distinct proliferation rates. Aside from being an interesting finding in its own right, this also validates the decision to select these populations based on their flow cytometry profiles, since it shows this phenotypic distinction is corroborated by distinct biological characteristics.

Confidence and tumour-specific effects

However, the confidence intervals on the long-term time-course values look very high. This is partly because towards the right hand side of the graph the logarithmic scaling can amplify quite small measurement uncertainties, and this noise has the potential to skew the regressions. However, it is also exacerbated by a second problem. Consider an ideal hypothetical sample in which the $EpCAM^{+}$ subpopulation divides twice as fast as the $EpCAM^{DIM}$ (both stochastically), and in which there is zero measurement error. This would generate two perfect straight lines on a transformed time-course plot. Now consider that for one mouse in the cohort, vascularisation of its PDX tumours is extremely poor. These tumours would grow more slowly overall. Even assuming that the perfect 2:1 growth rate ratio is maintained, both subpopulations would be incorporating EdU more slowly in this mouse, effectively shifting this time-point upwards on the graph (exemplified by the top-right panel of Figure 5.8) The

growth rate ratio ($EpCAM^+$ vs. $EpCAM^{DIM}$) has been perfectly maintained, but now the formerly variance-free regressions suddenly acquire variance. If this effect is pronounced enough, it may introduce sufficient variation that the subpopulations are no longer statistically distinct, even though the biology in this hypothetical example has remained perfect throughout.

In order to disentangle real biological effects from these tumour-specific effects, a different approach to analysis was taken. Two distinct subpopulations from the same tumour represent paired samples, however the previous analysis approach discarded this information, effectively treating them as entirely independent samples. To measure variation in a paired fashion, with this information retained, logarithmic difference was taken:

$$\ln(\,\%\text{ unlabelled } EpCAM^{DIM}\,) - \ln(\,\%\text{ unlabelled } EpCAM^{+}\,)$$

This value was then plotted over time. This is effectively measuring the degree of divergence of the two linear regression lines ($EpCAM^+$ vs. $EpCAM^{DIM}$) seen in Figure 5.8, minus any tumour-specific effects. By calculating difference on a tumour-by-tumour basis (effectively measuring the marking of one population relative to the other), tumour-specific effects can be negated, maximising the statistical power of the analysis. This comes at the cost of absolute measures; this approach therefore cannot provide DT values, but can demonstrate a trend over time with higher power. A regression trend and 2 SE confidence intervals can be calculated for this, and a p-value determined for the significance of the hypothesis that the $EpCAM^+$ subpopulation is proliferating significantly faster than the $EpCAM^{DIM}$ subpopulation.

As shown in Figure 5.9, 5 of the 6 time-courses show positive gradient, supportive of this conclusion, and 3 of these show significance to the $p \leq 0.05$ level (in this case, the probability that a non-zero gradient has been obtained by chance alone in the absence of a genuine correlation). This provides additional statistical power to support the hypothesis that these populations have distinct rates of proliferation, and suggests that this trend is present in most cases. However, the data from the OV04-496 cohort raises the possibility that in some circumstances, either sample-specific (e.g. treatment history) or cohort-specic (e.g. particularly harsh implantation conditions in cohort setup), this distinction may be lost or minimised to the point of irrelevance.

Short-term assays

However, the most important conclusion from these data-sets comes from the comparison between these long-term EdU traces, and the results of the short-term EdU pulse-chase experiments (as described in Figure 5.3). This short-term assay measures solely proliferation rate (through measuring percentage of cells in S-phase) as dividing cells are only given time to take up EdU, and are not given the chance to subsequently enter M-phase and divide. In

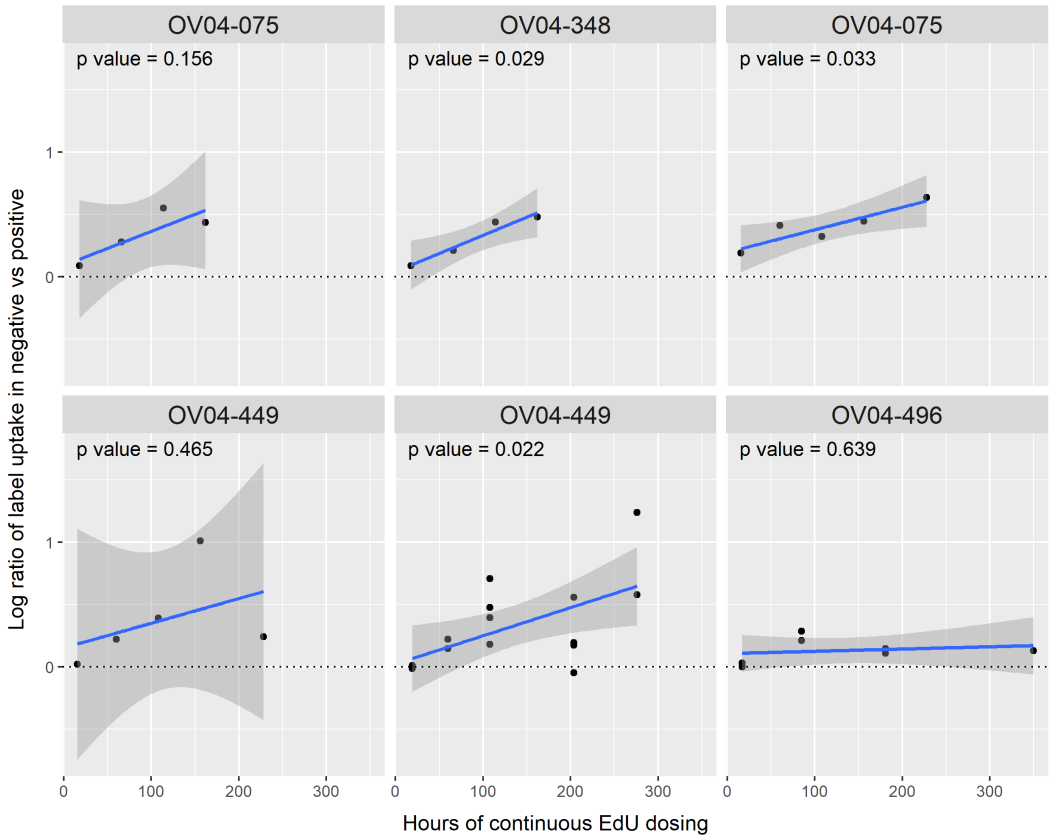


Figure 5.9: Logarithmic difference versus time for each of the long-term time-courses shown previously. Logarithmic difference is calculated by subtracting the logarithms of the percentage of EdU unlabelled EpCAM^{DIM} cells and the percentage of EdU unlabelled EpCAM^{+} cells. This effectively measures divergence in the two trends, and eliminates tumour-specific variation.

contrast, the long-term assay measures the destination of newly generated cells (by marking all cells involved in division since labelling began). Thus a classic cancer stem cell model would predict the $EpCAM^{DIM}$ population would be strongly marked in the short-term assay (because this is where proliferation is occurring), while being much more weakly marked in the long-term assay because the majority of the progeny generated would be $EpCAM^{+}$ cells. Therefore, the majority of EdU marking over the long-term would actually accumulate in a different subpopulation.

Short-term EdU data			
Type	Sample	EdU^{+} ($EpCAM^{+}$)	EdU^{+} ($EpCAM^{DIM}$)
Ascites	348	5.5%	2.0%
Ascites	072	6.0%	3.0%
Ascites	297	15.5%	6.5%
Ascites	366	6.5%	2.5%
Ascites	366	6.5%	3.5%
Ascites	366	12.0%	7.5%
Ascites	366	7.5%	4.5%
Cell-line	PEO-4	4.0%	2.0%

Table 5.1: Data from an initial series of short-term EdU assays. These samples were isolated PDXs and not part of experimental cohorts. Values given to the nearest 0.5%.

Short-term EdU trace data performed on a range of samples is shown in Table 5.1 and summarised in Figure 5.10. As can be seen from these samples, a similar pattern of EdU distribution is seen in these short-term experiments as was observed in the long-term time-courses, with the $EpCAM^{+}$ subpopulation accumulating EdU at between 1.6 x and 2.75 x the rate of the $EpCAM^{DIM}$ subpopulation (see right-hand panel of this figure). While there is, as expected, some variation in this ratio between experiments, there is a very clear trend in the comparative EdU uptake of the two populations ($p = 0.0012$). Indeed the variability in this ratio is perhaps surprisingly low given the large variation apparent in the absolute values. This would appear to strengthen the hypothesis that there is a distinct relation between the proliferation rates of the two subpopulations, as it appears conserved even when the absolute rate of tumour growth varies drastically. Another interesting observation is that the four examples from the same biological sample show no particular similarity, suggesting that this may be a loosely defined relationship and that factors extrinsic to the tumour itself (such as vascularisation) may have an influence over the relative proliferation rates of each subpopulation.

Paired short-term and long-term data

In addition to these independent short-term EdU incorporation experiments, a number of short term traces were performed on xenografts from the cohorts discussed previously. For

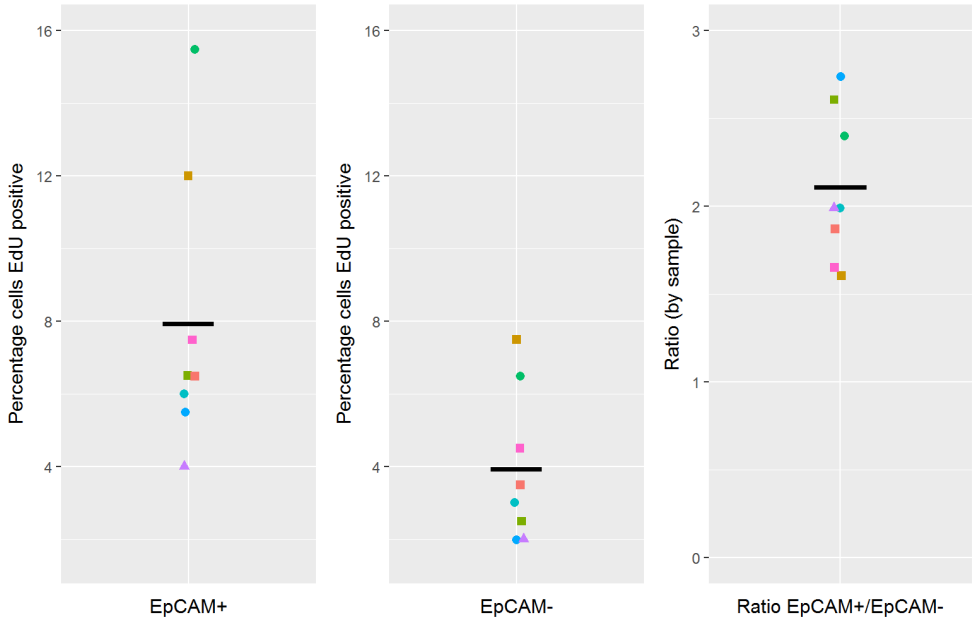


Figure 5.10: Summary of the data shown in Table 5.1. Each coloured point represents the same single experiment. Square points represent technical replicates of the same original patient tumour (0V04-366); triangles represents the single cell-line sample.

both sets of experiments, median cell division times (DTs) can be estimated from short-term data in a simple arithmetic fashion. The percentage of cells in a subpopulation marked with EdU, represents the percentage of cells in S-phase during the marking window. As per Figure 5.3, this marking window represents approximately the length of one S-phase. If 5% of cells are marked during this window, then 10 successive such periods are required to reach 50% marking. Thus DTs can be calculated by the equation below:

$$\frac{50\% \times \text{S-phase time}}{\%\ \text{EdU marked}} = \text{median division time}$$

This experiment means that accurate DTs are contingent on a good estimate of S-phase length. An initial estimate of 12 h was used for an initial analysis. A range of parameters, including S-phase length, from various tumours are collated in a review by Rew and Wilson (Rew & Wilson, 2000). This lists several reports from ovarian disease (in which disease subtypes do not appear to have been stratified (Erba et al., 1994)), all of which report S-phase within the 12-15 h bracket, with metastatic disease (more likely to be HGSOC) towards the lower end of this. As such, the initial estimate is likely to be accurate, if fractionally on the low side, and this 12 h value is used from here on.

For these cases where short-term and long-term data can be directly paired, the data are shown in Figure 5.11.

There is less of this paired data than would be optimal as the initial pilot time-courses

had too few mice for them to be spared for short-term assays, while the later time-courses (in which this was rectified) were mostly found to have the triple-subpopulation phenotype (which will be discussed separately). Random attrition of time-course mice due to health concerns also prevented collection of this data for one time-course , and reduced the power on others. While suboptimal, there is limited room for improvement in the approach used. When these xenografts were established, technical challenges with the EdU tracing protocol were still being overcome, and without any guarantee of a working experiment to use them in, it was not possible to justify setting up full-sized time-courses.

For the first two time-courses where paired data is available, the short-term and long-term data mirror each other well, and the relative relations are very similar. Differences in absolute value likely arise as a result of imperfect estimation of S-phase length. This is with the exception of OV04-496, where the data are too erratic to draw sound conclusions. This is the same dataset which previously broke with the pattern in the relationship between $EpCAM^{+}$ and $EpCAM^{DIM}$ subpopulations, so this may represent something aberrant about this dataset as a whole, rather than with the long-term / short-term correlation. The patient presented with very early disease limited to the fallopian tube (1C) but rapidly progressed despite adjuvant carboplatin chemotherapy and has subsequently had a very atypical pattern of lymph node and subcutaneous metastatic disease; thus this may simply represent an atypical biological sample. Alternatively, this replicate was particularly affected by the issue with PI fluorescence leakage so the issue may have been technical. However, since PI leakage causes mouse cells to shift upwards so they merge into the bottom of the $EpCAM^{DIM}$ subpopulation, a crude test for this issue is to restrict the gate to the top half of the $EpCAM^{DIM}$ region and see if this ultra-conservative gating produces a more expected result. In this case it does not resolve the anomalous results. I suspect the most likely explanation for this aberrant dataset is an atypical sample with erratic properties combined with a low-passage time-course. This time-course was only a second-generation PDX; in the context of a particularly atypically sample, the first-generation xenograft may not have grown entirely uniformly, and this may have led to a series of second-generation tumours which were not necessarily fully comparable. It is possible that this could be overcome with additional passaging, but this comes with its own drawbacks. Ultimately it is probably the case that particularly unrepresentative samples simply have to be accepted as such.

The prevailing concordance between short- and long-term data within each cohort, in addition to the same patterns being seen across short-term and long-term data more generally, is consistent with the results that would be expected from a purely stochastic pattern of division in which both subpopulations are self-supporting. A simple CSC model without transit-amplifying intermediaries would not be capable of producing these results, and a CSC model using TA cells would be highly unlikely to do so. It is clear from this work that both subpopulations appear to be being generated in a stochastic manner from their own compartment.

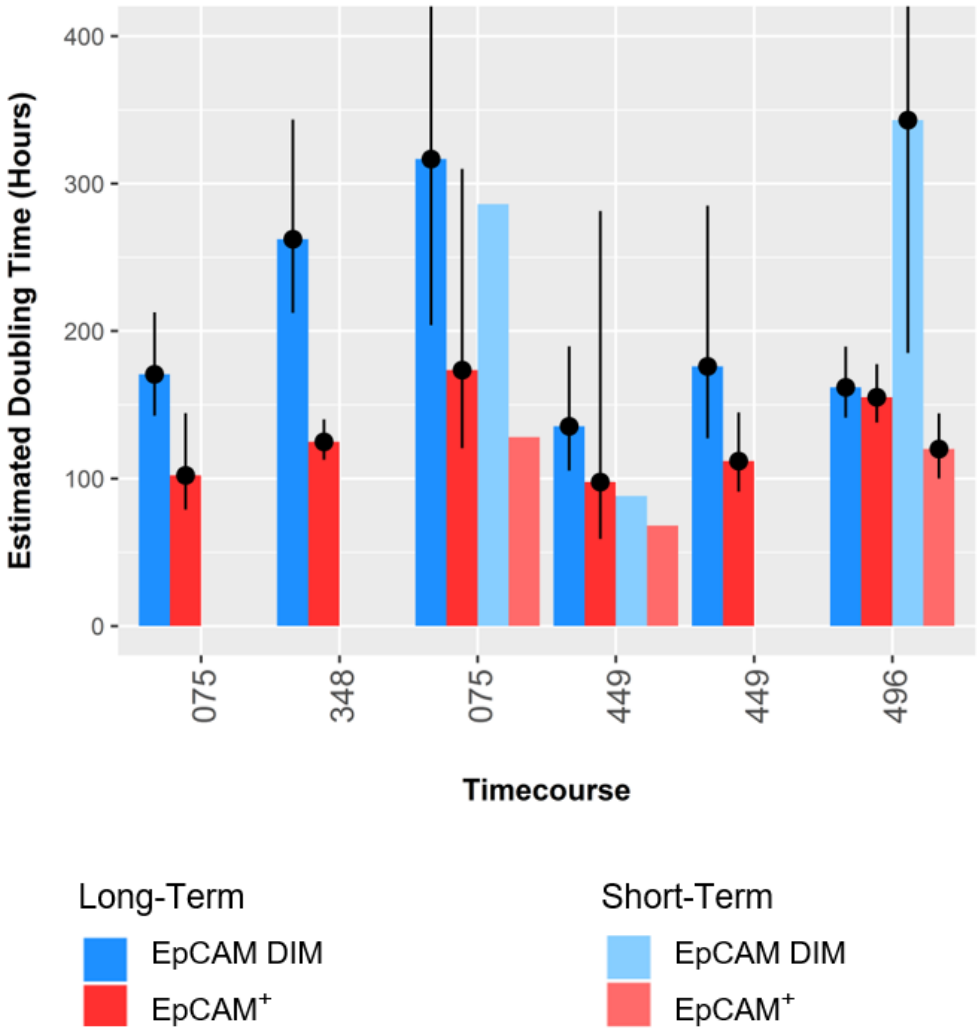


Figure 5.11: Median division times estimated from long-term and short-term data. Paired short-term data was not collected for the first two time-courses, and is missing for the second OV04-449 time-course due to mouse attrition; for the short-term data shown, n= 1, 1, and 2. Error-bars show 2 S.E. confidence intervals in each case as calculated from $EpCAM^{+}$ and $EpCAM^{DIM}$ treating data from each experiment in an unpaired fashion. *This is an inappropriate statistical test which drastically amplifies uncertainty, but is the only way to preserve absolute values suitable for plotting on this graph without losing statistical rigour. An assessment of uncertainty is made in a statistically appropriate and rigorous way in Figure 5.9 but these values cannot be plotted on a bar graph without significant and non-transparent extrapolation.*

Hypothesis modelling

While it is unlikely that a transit-amplified CSC model would generate data consistent with stochastic division, it is theoretically possible that under certain limited circumstances this type of division could give rise to data consistent with stochastic division. Simple hypothetical models were generated to investigate this further, and provide further evidence against the hypothesis of a CSC model generating a pseudo-stochastic pattern by chance.

Stochastic processes, by their very nature, are conveniently easy to model due to the simple mathematics that underlie them. For this example, an arbitrary starting number of 500 cells was selected. Each day, a proportion of cells divides (initially set to an arbitrary 5%), and is lost from the remaining unlabelled fraction (calculated as $500 \times (1\text{-}0.05)^{\text{Days}}$). Each dividing cell results in two daughters, so on the first day when 25 cells are lost from the unlabelled fraction, 50 cells will appear in the marked one. Initially the marked fraction was calculated by just doubling the number of cells being lost from the unlabelled fraction, with the total cell number being the two combined. However, this does not account for subsequent division within the labelled fraction which will also cause this population to grow. The solution is to calculate the number of unlabelled cells as above, and the total cell number by applying the same 5% expansion rate to the whole population (since all cells will divide regardless of whether they are labelled or not). Total cell number is thus calculated by $500 \times (1{+}0.05)^{\text{Days}}$. The number of labelled cells must be the difference in the two. Values were generated for days 0 to 10, along with the resulting percentage of labelled cells.

When plotted, these values give the graph in the top left panel of Figure 5.12. As can be seen here, these values give the characteristic stochastic curves seen in the experimental data (prior to log transformation). Since all the calculations outlined above depend on proportions, the number of starting cells is irrelevant; the graph would look identical if the input was 500 or 5 million (subject to integer rounding error). The only variable that changes the division profile is the rate of proliferation (the percentage of cell dividing per day, currently set to 5%). The resulting curve will always retain the same characteristic shape, but increasing this value will make the curve decay more steeply.

The value of 5% / day was an arbitrary approximation, but with the extensive experimental data on proliferation rate already presented, a value can be determined empirically and used to optimise the model. Combining all of the 6 time-courses discussed earlier and performing linear regression on the logarithmically transformed data (coerced to ln(100) at $t = 0$, and subsetted by subpopulation) gives two regression lines (EpCAM^{+} and $\text{EpCAM}^{\text{DIM}}$). These can be combined into a single linear regression, weighted by the relative size of each population (70% EpCAM^{+}, 30% $\text{EpCAM}^{\text{DIM}}$). This regression predicts 50% marking of the total population after 114 h. The input to the theoretical stochastic decay model can then be adjusted until it produces a DT of 114 h. A proliferation rate of 7.3% / day produces data which mirrors well the rate seen in the empirical experiment. This finalised model data is shown in Table 5.2, and is plotted in purple in the top-right panel of Figure 5.12.

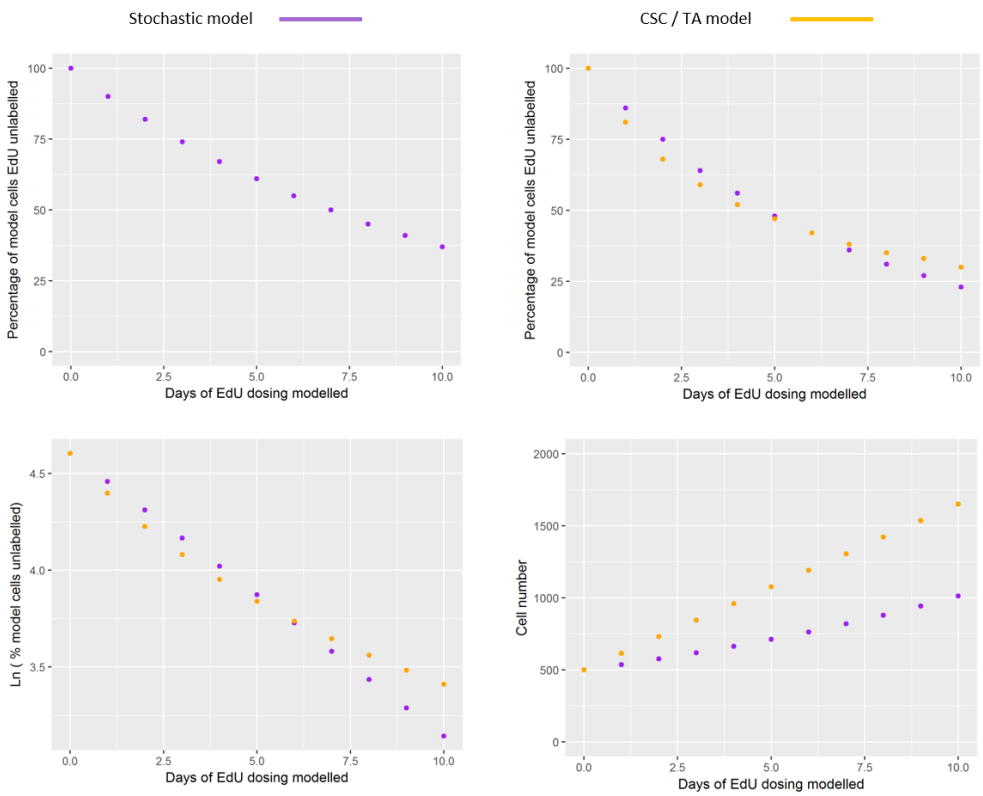


Figure 5.12: Decay rates of hypothetical division models. Top-left: a simple stochastic decay model; top-right: stochastic model (purple) adjusted based on empirical data, along with an attempt to generate similar data using a CSC/TA model (orange); bottom-left: the same data log-transformed; bottom-right: total cell number over time for these two models.

Stochastic Model					
Days	Unlabelled	Labelled	Total	% Unlabelled	ln(% Unlab.)
0	500	0	500	100	4.605
1	464	73	536	86	4.459
2	430	146	576	75	4.313
3	398	219	618	64	4.166
4	369	294	663	56	4.020
5	342	369	711	48	3.874
6	317	446	763	42	3.728
7	294	525	819	36	3.581
8	273	606	879	31	3.435
9	253	690	943	27	3.289
10	234	777	1012	23	3.143

Table 5.2: Hypothetical stochastic model with a marking rate of 7.3% of cells per day. The total population increases by this proportion each day, and cells are lost from the unlabelled population at this rate.

In order to determine the feasibility of the empirical data arising as a result of a transit-amplified (TA) CSC model, a second theoretical model was generated based on this hypothesis. This model starts with the same 500 cells. In this case, the vast majority of the initial cells are quiescent and will not subsequently divide. As such, this initial population remains unlabelled for the rest of the experiment. CSCs generate TAs which divide several times, giving rise to quiescent cells, before they themselves become quiescent. This results in a constant 'conveyor belt' effect, in which a small number of slow-cycling CSCs constantly replenish a slightly larger pool of TAs which in turn are constantly cycling to give rise to tumour replenishment and growth. The key feature of this system is that a fixed pool of CSCs result in a constantly replenished TA pool of fixed size, and therefore the generation of a constant stream of cells at a predictable rate. In reality, this model is a little inaccurate as it does not account for an increase in the number of CSCs as the tumour grows – this is actually a conservative representation, and a more accurate model that accounts for this would be even more extreme. Since all of these new cells are generated by division, they will be marked. Therefore, this constant stream of new cells can be modelled by simply adding a fixed number of new marked cells each day. From these two inputs, total cell number, percentage labelled, and the logarithmically transformed value can be calculated. There are two inputs this time, an arbitrary number of starting cells and an arbitrary number of cells added each day. When this data is plotted in the same way as with the stochastic model, the use of a proportional measure (percentage of cells labelled) means that the absolute numbers are once again not important. The only value which can influence the shape of the graph is the number of cells being added, relative to the starting number of cells (500 starting cells with 100 added per day would look identical to 1000 starting cells with 200 added per day). Trial and iteration of different input values was used to try and obtain a profile as similar as possible to that generated by the stochastic model based on experimental data. An input of 130 cells / day (26% of starting value; data shown in Table 5.3) gives the best approximation of the stochastic curve, and is overlaid in yellow in the top-right panel of Figure 5.12.

Transit-Amplified CSC Model					
Days	Unlabelled	Labelled	Total	% Unlabelled	ln(% Unlabelled)
0	500	0	500	100	4.605
1	500	130	630	79	4.374
2	500	260	760	66	4.187
3	500	390	890	56	4.029
4	500	520	1020	49	3.892
5	500	650	1150	43	3.772
6	500	780	1280	39	3.665
7	500	910	1410	35	3.568
8	500	1040	1540	32	3.480
9	500	1170	1670	30	3.399
10	500	1300	1800	28	3.324

Table 5.3: Hypothetical transit-amplified CSC model. The original unlabelled cell population is assumed to be negligibly proliferative; a tiny number of CSCs drive proliferation using transit-amplifying cells. The details of this are not examined, instead a fixed number of these (130) are assumed to be generated each day from this 'conveyor belt' process.

These profiles appear very similar, both on this graph, and on the corresponding logarithmically transformed plot in the bottom-left of Figure 5.12. While it is possible to distinguish them in these plots due to the stochastic model's idiosyncratic curve (which becomes linear in the ln-transformed plot), in the context of an empirical experiment with background noise and detection error, these plots would be indistinguishable. Ostensibly this may seem as if it undermines the assertion that the empirical data supports a stochastic model, since an alternative model appears to be capable of generating almost the same results. However, this dataset highlights the extremely artifical parameters that must be input into any CSC/TA model in order to replicate the empirical data. By quantifying the parameters necessary for a CSC/TA hypothesis to generate the empirical data collected earlier, the CSC/TA hypothesis can be falsified by demonstrating that these parameters are not present in experimental circumstances.

In particular, what makes the CSC/TA model strikingly artificial is its rate of proliferation. The stochastic model allows a continuous *rate* of movement of cells from the unmarked to the marked population, resulting in a smooth increase in marking (and thus in the percentage of marked cells) regardless of population size. In a CSC/TA model, the vast bulk of the initial population remains unmarked, and the only way for this model to recapitulate the same smooth increase in the proportion of marked cell is to generate new cells by rapid proliferation. To maintain a steady drop in the percentage of cells unlabelled, this model therefore requires a precipitous increase in cell number over time, as shown in the bottom-right of Figure 5.12.

Thus by measuring the growth of tumours empirically, this hypothesis can be falsified by demonstrating that the required proliferation of the entire tumour is not observed.

Tumours were routinely measured using callipers (length, width and height), and the size

in mm estimated from $0.5 \times length \times width \times height$. Tumours were measured accurate to the nearest millimetre, and values below 3 mm were neglected. Since tumours were mostly grown by implantation of solid material, most had 'palpable tumours' from implantation. In some cases, these would gradually shrink, becoming undetectable, before a tumour subsequently grew from the same site. As this experiment (and indeed the entire project) is interested in the growth of established tumours, not in tumour initiation and establishment, all measurements prior to the point sustained growth was detected were discarded; this timepoint was then treated as day 0. Data from 65 tumours was collated and plotted (left panel, Figure 5.13) on logarithmically transformed axes. Linear regression through this dataset allowed an estimate of the daily growth rate (the percentage a tumour would grow by per day). This was calculated to be 5.78% (95% CIs 4.51% – 7.07%).

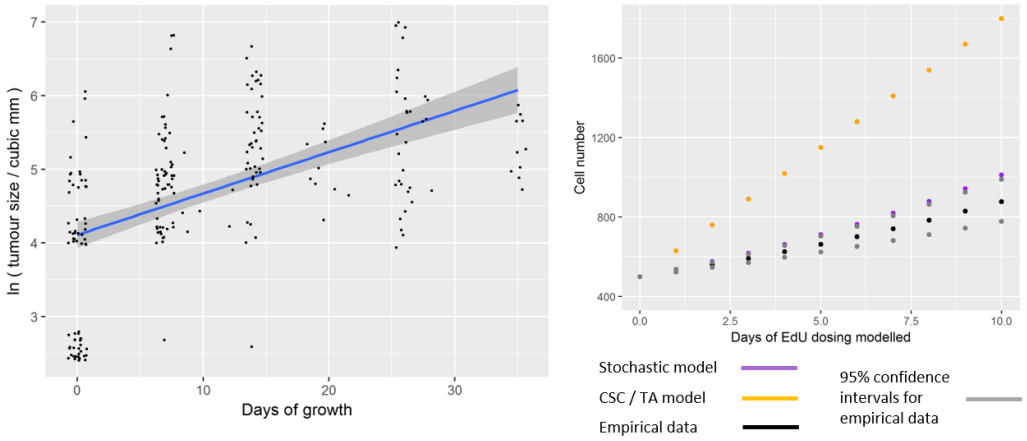


Figure 5.13: Left panel shows collated empirical tumour growth data. Size (in mm) is taken from tumour calliper measurements of tumours, and is shown logarithmically transformed. Rate of proliferation is estimated by linear modelling. Right panel shows how the empirical growth rate compares against the two model datasets shown in Figure 5.12. The black curve shows the estimate from real data, and the grey curves show 95% confidence intervals around this.

As can be seen from Figure 5.13, the empirical data map fairly well to the stochastic model data, falling only just beyond the upper 95% confidence interval of the real data. The CSC model meanwhile produces results which lie ∼10 standard errors off the experimental data.

In conclusion, this assessment shows that while it is possible for a CSC model to produce a good approximation of the empirical EdU uptake time-courses shown, it critically cannot also produce the empirically observed tumour growth curves in the same experiment. By contrast a stochastic model with parameters such that it mirrors the empirical EdU uptake curve, automatically results in a growth curve which appears similar to that observed empirically in real xenografts. This analysis of the available data would appear to preclude not just a simple CSC model, but also all variants of a CSC/TA model, regardless of the phenotype

or growth parameters that any TAs might have. This provides a strong argument in favour of a stochastic model of HGSOC proliferation.

Apoptosis in HGSOC and impact on assays

One possible weakness of this modelling is that it does not account for apoptosis. It is possible that extensive apoptosis could remove cells at a sufficient rate to make a CSC model viable again. Measuring relative rates of apoptosis is a very straightforward process involving immunohistochemistry staining for cleaved-caspase-3 (CC3).

To do this, histology sections were cut from 4 xenograft samples and stained by IHC by the CRUK Cambridge Institute Histology Core. These sections were then imaged using an Axio Scan.Z1 imaging platform (Zeiss) and analysed using HALO software (Indica Labs). This software takes the image, identifies cells based on the IHC counter-stain, and then assigns cells as either $CC3^+$ or $CC3^-$ based on the primary IHC staining. An example of this staining is shown in the top-left of Figure 5.14, and a summary of the resulting staining rates shown in the bottom-left. On average, 0.7% of cells are estimated to be $CC3^+$.

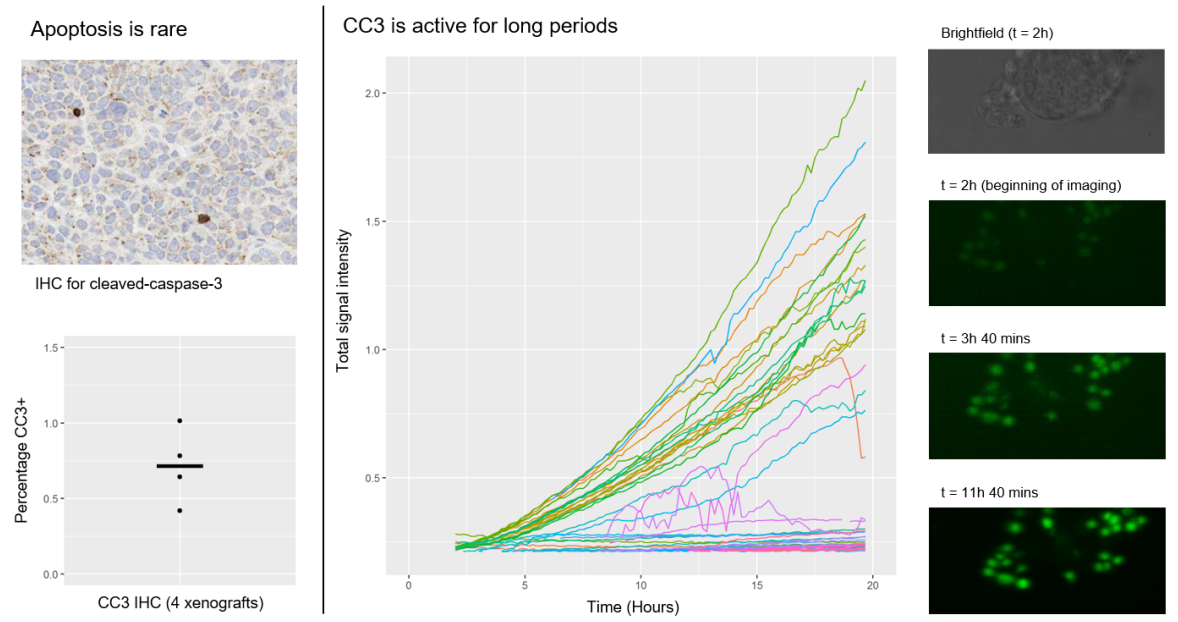


Figure 5.14: Apoptosis is minimal in xenografts. Top-left: typical IHC staining for cleaved-csapase-3; bottom left: quantification of $CC3^+$ rates in 4 xenografts (OV04-366, -581, -297, -338); centre: signal (total intensity from 48 cells) from fluorescent CC3-substrate over the course of an *in vitro* organoid assay; right: brightfield of the organoid, and sample fluorescence images (collapsed z-stack) from this time-course.

However, for the purposes of this experiment, this needs to be translated into an absolute rate of apoptosis (percentage of cells undergoing apoptosis per day). In order to do this, the length of time an apoptosing cell should be CC3-positive for needs to be established.

The details of apoptotic systems have been investigated in some detail, and the CC3 cascade is shown to be completed in an individual cell in a matter of minutes, with cell populations treated with high doses of various toxic reagents entering apoptosis over <6 h (Rehm et al., 2002). However, once CC3 has triggered downstream effects, it has no further role in the process of apoptosis – as a result there has been little attempt to investigate its longevity.

This indicates that CC3 positivity should appear rapidly in the CC3 IHC assay, but does not indicate how long that positive signal will remain for. If cells remain $CC3^+$ for a few hours, then 0.7% of cells being $CC3^+$ at a single point in time would translate into a relatively high rate of absolute apoptosis. For example, if cells are $CC3^+$ for only 3 h then 0.7% would translate into an absolute apoptosis rate of 5.6% day (24 h/3 h x 0.7), bringing the CSC theory back into contention. Thus it is important to determine how long apoptosing cells are $CC3^+$ for.

Tawa *et al.* investigate CC3 degradation in more detail (Tawa et al., 2004). They performed western blot assays for both caspase-3 (pro-enzyme) and CC3 (the activated form) for cells before and after activation and noted that there was an apparent loss of protein in the cells where caspase-3 was activated, which they attribute to degradation. Inhibition of CC3 activity using an active-site inhibitor abrogated this loss, as did point-mutation of key catalytic residues, leading the authors to conclude that CC3 activity is involved in its own elimination. This paper lacks clear timescales (since that was not the purpose of this study). However, from the data shown (Fig. 1) it appears that after 4 h there is significant degradation of CC3 (compare the quantity of protein in the uncleaved caspase-3 [top band, left] to the vastly reduced amount of CC3 observed 4 h later [bottom band, right] – the latter would be expected to be roughly half the strength given the relative size of the two proteins, but is actually many orders of magnitude weaker, indicating significant degradation).

To test this empirically, the activity of CC3 needs to be assessed on an extended timescale at the single cell level. Since this experiment is not practical *in vivo*, the next best option is a 3D *in vivo* model.

Organoid models have been developed within the Brenton Laboratory by Dr Debbie Sanders from multiple HGSOC tumours and passaged xenograft models. In order to investigate apoptosis within these, they were subject to live-imaging on a Lightsheet Z.1 (Zeiss) operated by the CRUK Cambridge Institute Microscopy Core. The cells were grown in plastic capillary tubes in growth-factor enriched media for two days, then treated with campothecin (to promote apoptosis) and a fluorescent CC3 substrate (Essen Bioscience) to assay CC3 activity. Two hours later, the cells were installed in the instrument, suspended in 20 ml of the same treated media, which could slowly diffuse into capillary. Detailed methods are described in subsection 2.7.2. Thus the enzyme has immediate access to only a small amount

of substrate, but a slow but constant feed of it over the course of the experiment, allowing activity to be assessed over a long period.

The centre panel of Figure 5.14 shows the resulting fluorescence signal over time for 48 cells followed using this live-imaging. Organoids derived from patient OV04-348 were used for this experiment. Of these 48 cells, 26 show only minimal intensity throughout the experiment, and are likely autofluorescent bodies. The remaining 22 all appear to show a sustained increase in brightness over the course of >15 h. The right-hand panels show a brightfield image of the major organoid assayed, along with fluorescence images (collapsed z-stacks) taken over the course of the experiment.

This appears to suggest that although CC3 exerts its biological effects in a matter of minutes, apoptotic cells retain CC3 activity over a much longer timeframe, indicating that the protein is still present in cells many hours later. This indicates that the CC3 IHC assay is able to capture cells within a broad time-window within the apoptotic process.

Given a minimum for this window of 15 h it is possible to estimate a maximum rate of apoptosis / day using the following formula:

$$\frac{\%\text{CC3}^{+}\text{ (IHC)}}{\text{time-window in which cell is CC3}^{+}} = \text{rate of apoptosis per day}$$

The IHC results presented previously give a maximum of $\sim$1% of cells CC3^{+}, and an average of $\sim$0.7%. Assuming 15 h for apoptosis, this gives a maximum daily rate of 1.6%, and an average of 1.1%.

Even this maximum rate of apoptosis would have litte effect on the previously described models, as seen in Figure 5.15. If anything, incorporating apoptosis simply makes the stochastic case more compelling. The CSC model continues to require precipitous growth that is not observed empirically.

The experiment measuring CC3 activity is imperfect in that it necessarily uses an *ex vivo* assay. However, it is not practical to conduct this experiment *in vivo*, and 3D organoids represent the gold-standard in *in vitro* methods. Another potential weakness of the experiment was the initial default signal analysis algorithm, which output intensity values in the form of the maximum signal density across the area of the cell (i.e. the brightest pixel in the assigned area). As this experiment studies apoptotic cells, which might be expected to condense and shrink over time, the density of signal might be expected to increase, even if the absolute signal remains the same. As such, this measure was somewhat unreliable. In order to improve on this, an alternative analysis algorithm was applied, where values were calculated by summation of signal across a cell. The Imaris system is efficient at identifying individual cells (by areas of dense fluorescence), but relatively poor at accurately delineating their edges, so instead the bounds were defined as a 5 µm circle around peak fluorescennce. In practice this captured the majority of each nucleus, and crucially avoided capturing fluo-

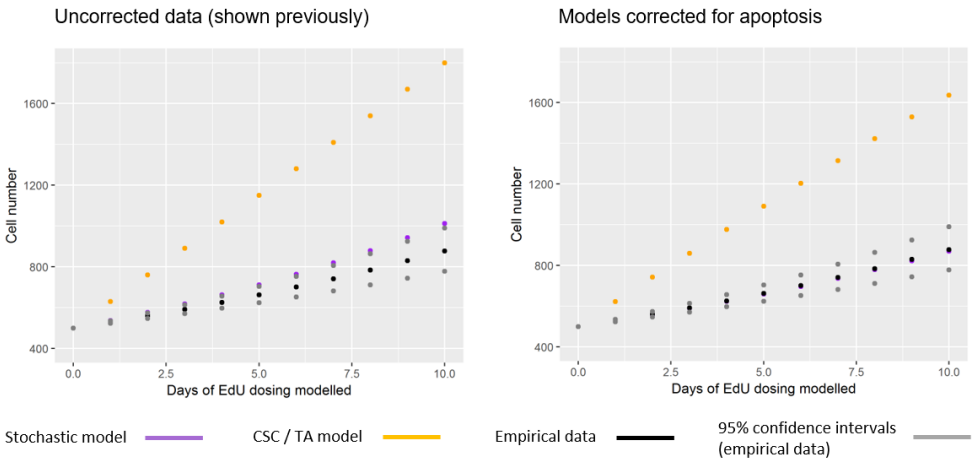


Figure 5.15: The previously shown hypothesis modelling figure (left) compared against data with apoptosis corrected for (right). In the latter case, the maximum rate of apoptosis observed in the empirical data (1.6% per day) was incorporated into both hypothesis models. Empirical data is unchanged. Stochastic model data (purple) is hard to see in the right-hand plot as it is directly overlaid on the empirical data (black).

rescence from neighbouring cells. While the edges of the nucleus were clipped, these showed much less fluorescence than the centre, and this clipped area was not seen to noticeably shrink over time. It is data from this improved method that is shown in Figure 5.14; this is broadly similar to the data from the first method, but with drastically less noise. This similarity indicates that the cells are not shrinking, while the more consistent trace implies that the improved method is significantly more reliable.

One final potential variable arises from the mechanism of the fluorescent substrate. This molecule is able to permeate the cell cytoplasm, where it remains in an inactive state, but is not able to penetrate the nucleus. Cleavage by CC3 releases the fluorescent component which is able to localise to the nucleus, where it is able to bind to DNA; upon DNA binding it becomes fluorescent. The nuclear localisation is likely to be rapid and thus not be a rate limiting step (although this cannot be entirely excluded). However, fluorescence being dependent on DNA binding means it is possible that signal may be influenced by DNA state. Since the state of DNA condensation changes significantly during apoptosis, there is a possibility that the signal intensity may be being influenced by epigenetic changes. Excluding this possibility would require conducting this experiment with a substrate which is activated in a non-DNA-dependent fashion. Since no such reagent is readily available, this experiment represents the best possible available evidence that CC3 is active for a prolonged period in apoptotic HGSOC cells.

Additional data was captured beyond the 20 h mark shown above. However, the system had to be reset between these two sets of captured data, making it harder to draw rigorous conclusions. The data is presented in Figure 5.16 with this caveat. The datasets were joined manually, and the two halves have been deliberately left unjoined to make this clear. An

hour's gap has been left as an approximation of the time the instrument was not recording for; colour-coding is not conserved across the boundary, although in many cases it is intuitive which cell is which.

This extended dataset is erratic (as might be expected given that cells are not alike), but towards 25-40 h mark the signal from a number of cells levels off and then drops to near-zero. The former presumably represents the loss of CC3, while the latter is presumably the beginnings of cell destruction (strictly it may just be DNA/nuclear destruction since this is where the marker is localised and active). Notably some other cells appear to continue accumulating signal, indicating ongoing CC3 activity, and a degree of heterogeneity in the apoptotic process. However, the combined data make it possible to draw a robust approximation of the length of CC3 activity within a cell – on the order of 40 h.

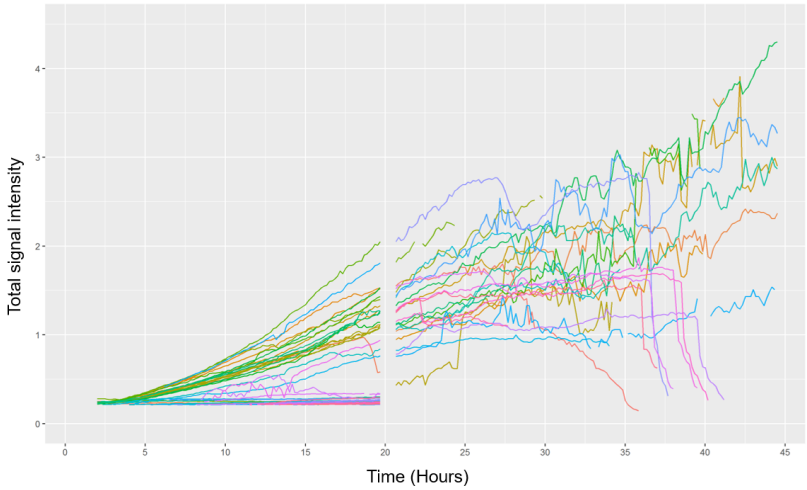


Figure 5.16: Extended plot of CC3 data from lightsheet experiment. Datasets were joined manually, so the join (and thus susbsequent timings) is improvised; colour-coding is not conserved between sets.

In the context of previous literature, it could be argued that even the conservative value of 15 h used for the length of apoptotic window in the previous calculations are somewhat high. Certainly some previous literature has shown notably lower values, e.g. the 3.5 h reported by Potten (Potten, 1996), however this very same dataset also reports a 12 h window in subtly different circumstances. Firstly, it is worth noting that these values represent a measurement of a different process: the half-life of wash-out of apoptotic bodies following a toxicity pulse rather than measurement of an average length of the apoptotic process in a cell - this will be discussed further shortly. The large discrepancy between the two experiments related by Potten also indicates that the length of the apoptosis detection window can vary significantly, and that context is crucially important. As such, the most reliable data will come from experiments conducted on HGSOC cells, which represent the right tissue and

transformation state.

It is also worth noting that the differential in the results reported by Potten appears to arise as a result of the induction method, with radiation inducing slow destruction and hydroxyurea inducing rapid destruction. The author attributes the difference to hydroxyurea being cell-cycle dependent, while radiation is much less so. The exact mechanisms of staurosporine (used in the experiment described in this thesis) remain somewhat cryptic, and although it appears to have significant interaction with the cell-cycle, its acute toxic effects appear to be somewhat independent (Qiao et al., 1996). This point may in any case be academic as staurosporine is being used to measure the length of the process which can then be used to inform analysis of the CC3 IHC experiments. The CC3 study assesses the extent of background apoptosis in PDX tissue, in which situation the exact cause of apoptosis is unknown. There may also be a range of cell-cell and bystander effects impacting apoptosis in all of the toxicity studies described which would not be expected in the original pathophysiological environment where only isolated cells are entering apoptosis, and their neighbours are unaffected. Finally, the Potten experiments appear to have been performed in murine models, versus the previously described experiment which was performed in HGSOC organoids. These approaches have different strengths and weaknesses, with the latter dependent on an *in vitro* system (albeit a gold-standard one), while the former approach is dependent on the assumption that a different tissue, transformation state, and even organism will have similar underlying apoptosis lengths. The literature already discussed suggests that rates are different between tissues. Equally, as previously discussed, the literature surrounding measurements of the length of S-phase indicate that this process appears to be notably longer in human tumours vs normal mouse tissue; it would not therefore not be surprising for a similar differential to be observed in apoptosis. As such, regardless of the method used in any experimental measurement of the apoptosis window, it is hard to be certain that this value remains appropriate in the original context of *in situ* PDX tumours, and it is a case of attempting to minimise the extent of contextual discrepancy between the model and *in situ* environment.

Returning to the values reported by Potten, these are not equivalent to value derived in the previously discussed experiment assessing apoptosis length in HGSOC organoids. Potten measured the half-life for the wash-out of apoptotic bodies after a toxic pulse, whereas extrapolating the extent of apoptosis via CC3 IHC as previously described requires the mean-lifetime of cells undergoing apoptosis. If 50% of apoptotic bodies were lost by 12 h, and 100% by double this (24 h), then the two would be equivalent. However, this is an exponential decay process, so by 24 h only 75% of apoptotic bodies would be lost. As such, the mean-lifetime (the length of time an 'average' cell takes to go through the apoptosis process), is higher than the half-life of decay. Thankfully the highly defined nature of exponential decay means that it is possible to easily convert a half-life for a process into a mean-lifetime (tau from classic exponential decay mathematics; half-life over $\ln(2)$). As such, the values of 3.5 h and 12 h can effectively be multiplied by a factor of 1.44 to give mean-lifetimes of 5.0 h and 17.3 h which can then be applied to the previous models.

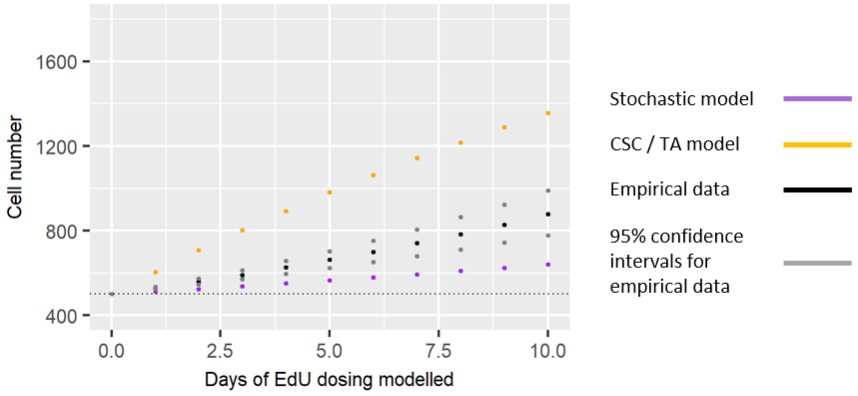


Figure 5.17: Test of robustness for the modelling approach; data favouring high apoptosis rates is cherry-picked from various experiments and applied out of context to see if even these extremes can bring the CSC model back into contention; as can be seen the stochastic model still remains more likely (4.8 standard errors versus 8.5). The stochastic model is shown in purple, the CSC model in yellow, the estimate of empirical growth from real data in black, and 95% confidence intervals around this empirical data in grey.

To test the robustness of the stochastic hypothesis set out in this chapter, the most unfavourable data from this assessment (the figure of 3.5 h for apoptosis half-life) was combined with the most unfavourable data (i.e. with the highest level of apoptosis) seen in the PDX work discussed in this chapter. Taking the mean-lifetime of 5.0 h, and combining it with the highest apoptosis index seen as part of the research discussed in this thesis (index of 1% as seen in the bottom-right of Figure 5.14), results in the hypothesis model shown in Figure 5.17. This model is highly equivocal, and it would be hard to argue that it supports the stochastic hypothesis, which now lies 4.8 standard errors off the (unchanged) empirical data. However, it is even clearer that, despite cherry-picking the most convenient data from different experiments, and assuming the compatibility of apoptotic index data and measurement time-window data from different experiments, biomarkers, tissues, transformation state, and ultimately species, the CSC model has not been brought back into contention. While under these artificial extreme conditions the stochastic model is a poor fit at 4.8 standard errors off the empirical data, it still remains a better model than the CSC model which still lies a full 8.5 standard errors off the empirical data.

It is also worth reiterating the construction of the CSC model is conservative in that it assumes the number of CSCs and TAs (and hence cell proliferation rate) is constant and that this pool does not grow with the tumour, even as the tumour expands. In practice, CSCs might be expected to proliferate (for example Doug Winton's laboratory have shown that intestinal crypts have 5-7 stem cells but that this rises to 8-9 in adenomas; (Kozar et al., 2013)), resulting in an exponential growth rate than is even more extreme than the one modelled here. As such in this model the apoptosis term is taking full advantage of multiplicative compounding while being offset by growth which is being modelled in a simpler linear fashion, even though the limited available evidence suggests this is probably

not representative (Kozar et al., 2013). The higher the apoptosis rate, the more impact this differential has, further skewing the data towards the CSC model. As such a more representative model would be even less supportive of the CSC hypothesis.

In addition, this model makes the assumption that the apoptosis rate is equivalent across CSCs, TAs, and bulk tumour cells. As such it makes the assumption that apoptosis does not impact cell labelling as it depletes all populations (CSCs, TAs, and bulk, both labelled and unlabelled) at the same rate, leaving labelling ratios unaffected, and meaning that the labelling curves shown in the top-right of Figure 5.12 can be assumed to still be accurate. In reality there is evidence to suggest that apoptosis is more pronounced in certain populations (Potten, 1977) (Potten, 1992), and that these cell populations are stem-like cells (Buczacki et al., 2013). As stem-like cells label more quickly (or indeed at all), higher apoptosis in these populations would mean the bulk tumour population is not being depleted as quickly by the process. This would in turn necessitate even higher proliferation in a CSC system than is being modelled here, meaning a more realistic modelling of the CSC hypothesis would be even less consistent with the empirical data than set out here. This further reinforces the alternative stochastic hypothesis.

Finally, the maximum yield of apoptotic bodies is seen 3 h to 6 h after exposure to cytotoxic agents, as measured by counting of apoptotic bodies (Potten, 1992). It seems unlikely that peak drug toxicity would occur in a delayed fashion, which suggests that there may be several hours between biochemical initiation of apoptosis (via caspase activation) and formation of an apoptotic body visible to crude histology. During this time, cleaved-caspases could be detected in dying cells, while crude histology would not identify them. As such the time-window of apoptosis by measured by cleaved caspases may be higher than that measured through crude histology. As such, taking an apoptotic index measured cleaved caspases, but using the time-window of the apoptotic process as measured by crude histology, likely overestimates the total apoptosis rate, creating an artificial over-inflation of the importance of the process.

In conclusion, even with such a plethora of artificial conditions tailored solely to coerce the CSC model into fitting the empirical observations, it remains effectively impossible to render it consistent with the empirical data. This strongly reinforces the case for an alternative model, one that the stochastic model meets elegantly and accurately when the most appropriate data is incorporated. With the exclusion of the possibility of significant apoptosis, it can be concluded – based on the modelling data previously presented – that neither a CSC-only, nor a CSC/TA model can satisfactorily explain both the nucleoside incorporation data *and* the tumour growth data of the HGSOC xenografts. It is thus reasonable to conclude that HGSOC is driven by stochastic mechanics and that cancer stem cells do not play a significant role in the growth of established HGSOC tumours.

Three-subpopulation analysis

In two xenograft time-courses, along with the single cell-line xenograft experiment, the third $EpCAM^+/CD90^-$ ($CD90^-$) tumour subpopulation was seen (as per Figure 4.8). A similar analysis was applied to these datasets.

All of the long-term time-course data shows $EpCAM^+$ proliferating slightly faster than $EpCAM^{DIM}$ (Figure 5.18), which in turn proliferates slightly faster than $CD90^-$. However, in all cases the differences appear small, and the variance appears too high to draw particularly reliable conclusions. From the initial transformed plot, the PEO-4 cell-line time-course appears to have drastically less separation between regression lines for the three subpopulations. While it might be instinctive to attribute this to the artificial nature of cell-lines and subsequent loss of distinct proliferation patterns, it may simply be because the cell-line is proliferating much more slowly. With a lower signal, the signal-to-noise ratio is much higher, and the variance is able to easily swamp the genuine differences in the populations. The DTs calculated from these data are plotted in (Figure 5.19), alongside the paired short-term data. The size of the error bars on the cell-line time-course are heavily influenced by the short measurement period and the signal-to-noise issue previously mentioned, although the statistic issues around quantifying uncertainty in form of absolute numbers (precluding treating data from the same sample in a paired fashion) also plays a major role. This time-course was an early validation experiment used to test the validity of the method, and is very underpowered, but is included here because it is interesting to note (at least in the case of PEO-4) that cell-line xenografts appear to produce broadly the same patterns as seen in PDX samples (at least when implanted *in vivo*). This would suggest that even cell-line xenografts might have some merit for studying phenotypic heterogeneity. While PDXs are likely to be more representative in general, for some basic research cell-line xenografts may offer some advantages due to their more easily distinguished phenotypic subpopulations.

As with the previous data, the analyses shown in Figure 5.19 include the influences of tumour-specific effects, and these can be removed in the same fashion in order to assess trend and statistical power. This type of statistical test is by its nature comparative, and compares one subpopulation against another. Previously the $EpCAM^{DIM}$ subpopulation was compared against the $EpCAM^+$ (logarithm of the former subtract the logarithm of the latter). For the purposes of assessing three subpopulations, the $EpCAM^+$ subpopulation was treated as the 'baseline', and both other subpopulations compared to it. $EpCAM^+$ was selected for two reasons. Firstly, because it is typically the largest subpopulation, and it is logical to compare the rarer subpopulations against the most common. Secondly, $EpCAM^+$ represents the fast-cycling, apparently chemosensitive population, which is of less interest within the context of the broader translational scope of the project, since these cells do not appear to be representative of minimal residual disease. The results of this analysis (Figure 5.20) show that while there is a positive trend for $EpCAM^{DIM}$ in all cases (suggesting that this subpopulation is again dividing more slowly that the $EpCAM^+$), the two subpopulations are not significantly distinct. For the $CD90^-$ subpopulation however, removing

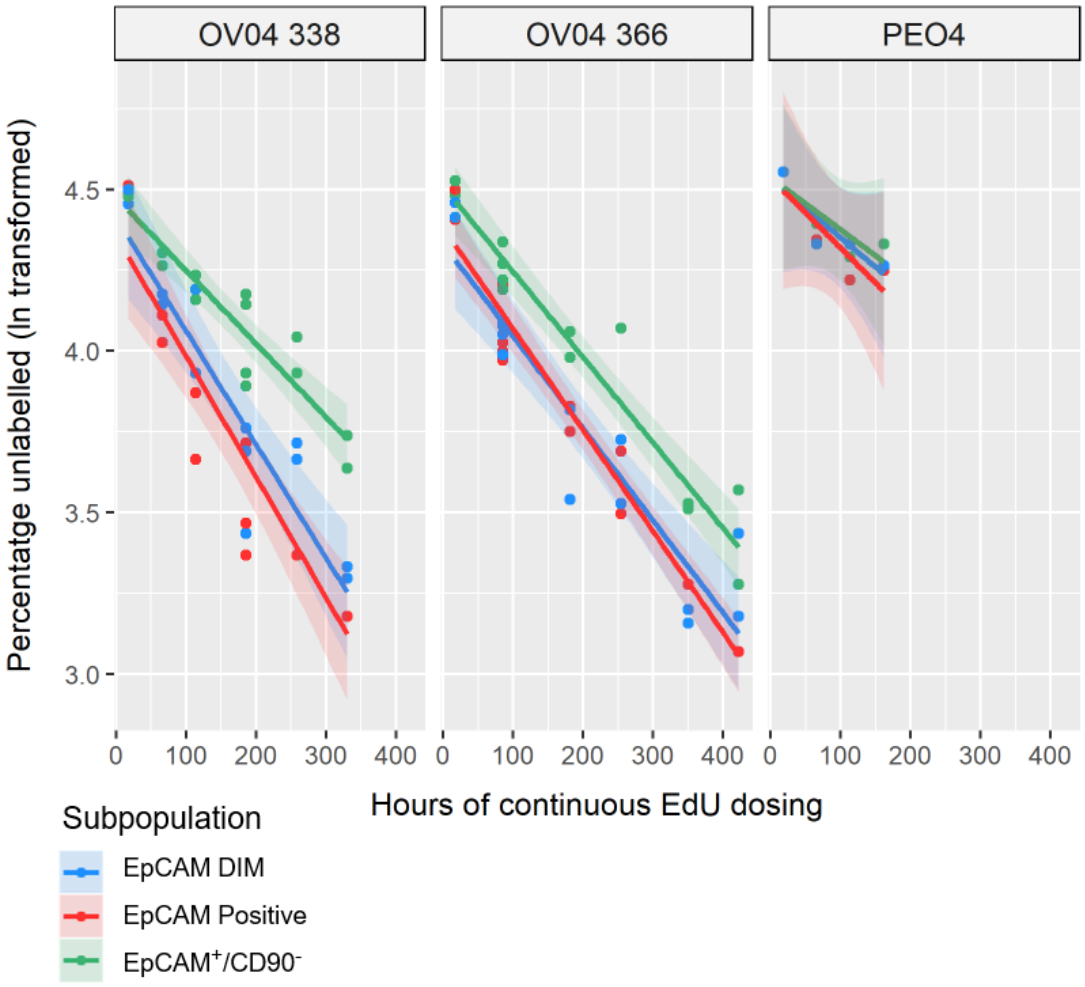


Figure 5.18: Logarithmically transformed plots for the two PDX and one cell-line time-courses in which the CD90$^-$ subpopulation was also seen. Regression models used to calculate subsequent DTs are also shown. Shaded areas show 2 SE confidence around the regression model.

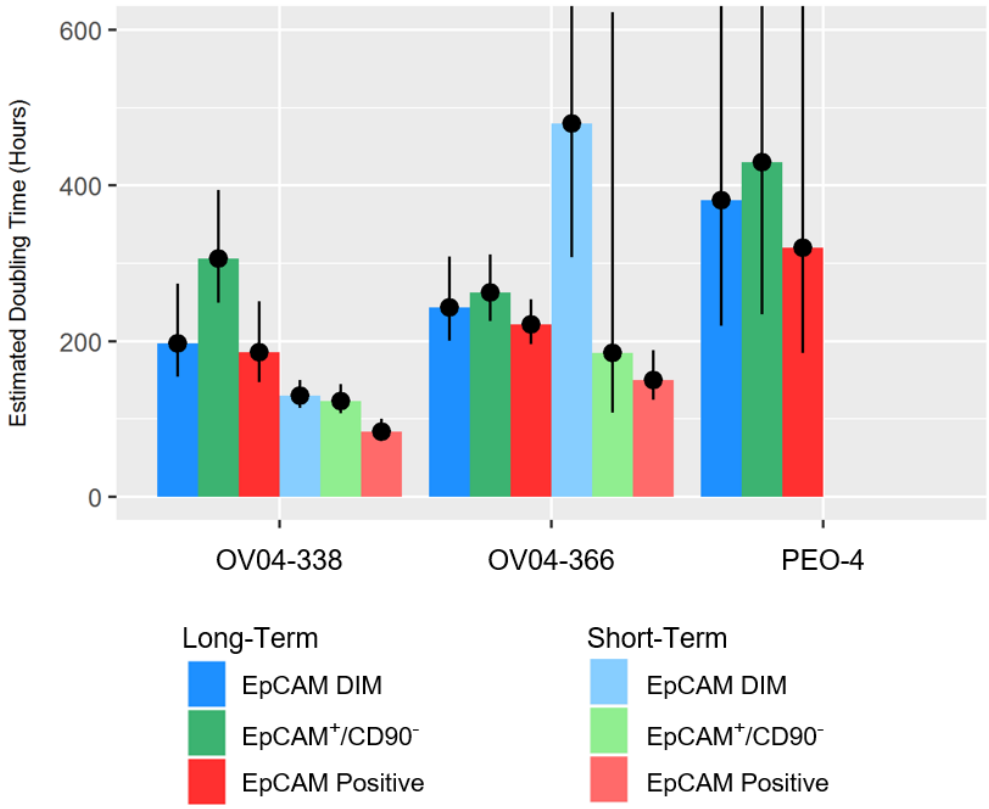


Figure 5.19: Median division times for time-courses with three tumour subpopulations, calculated from both long-term and short-term EdU assays. Short term data n = 2 for each time-course. Error-bars show DTs corresponding to 2 S.E. confidence intervals around the regression model.

tumour-specific effects shows a clear and significant positive trend in the PDX samples (indicating that this subpopulation is dividing significantly slower than the EpCAM^{+}). This trend is pronounced enough that it even approaches $p \leq 0.05$ in the cell-line xenograft, despite the short time-course and the low overall proliferation rate. These observations raise some interesting questions about the role of the CD90^{-} subpopulation, how it interacts with the other two, and why it is only apparent in some cases.

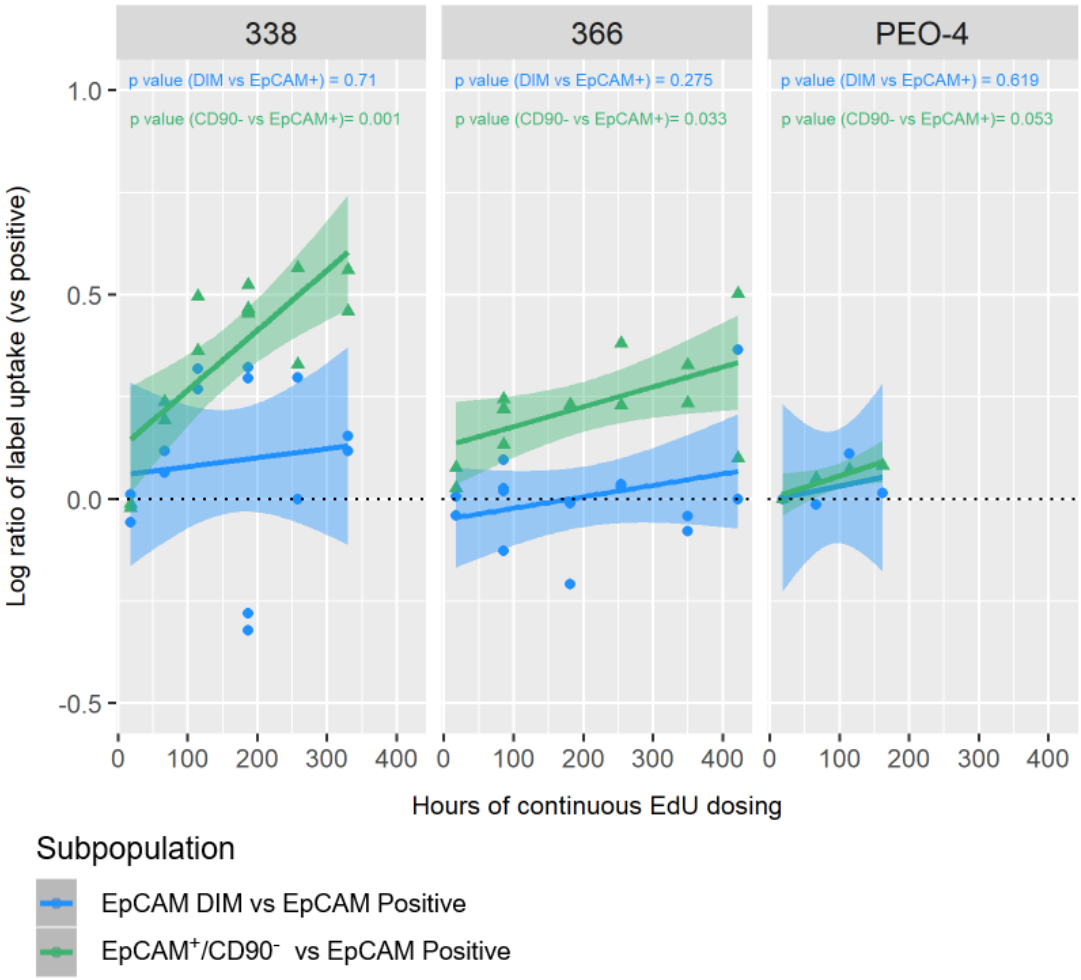


Figure 5.20: Relative proliferation rates, using the paired analysis to eliminate tumour-specific effects. EpCAM$^{\mathrm{DIM}}$ and CD90^{-} are compared against the EpCAM^{+} subpopulation.

5.3.3 Validation of possible confounding factors

All of this previous work is dependent on the assumption that proliferation, as measured through EdU uptake, is solely a function of time and tumour division mechanics. In order for these assumptions to be relied upon, it is important that the effects of potential confounding factors are considered and excluded.

Three potential confounding factors have been identified which have mechanistically plausible mechanisms through which they might distort the data previously presented. These include variation in tumour size, variation in tumour vascularisation, and the impact of EdU itself on proliferation rate. In particular, in each case there are also reasons to believe that these features will not be randomly distributed across datasets, and therefore if EdU uptake is influenced by these features, it may not be possible to reliably attribute previously observed trends to tumour division mechanics alone, rather than to the impact of these confounding variables. The following section considers each of these three factors in turn, detailing the rationale for the concerns, and then presenting an assessment of the available evidence.

Tumour size does not significantly impact proliferation and EdU labelling

The first concern is that variation in tumour size may be impacting EdU uptake measurements. If larger tumours incorporate EdU at a slower rate (for example due to the reagent having limited access to the core of the tumour), then proliferation readout (as measured by change in EdU labelling) would slow as the tumour grows larger, even if the true rate of unlabelled cells entering division remains constant. While the use of multiple animals via a time-course experiment means that there is not a direct relation between tumour size and time post-EdU-treatment (as there would be if a single sample was measured over a period of weeks, since growing tumours get consistently larger), it would be reasonable to expect that tumours towards the end of time-courses would on average be larger as they have had additional days or weeks to grow. The conclusion of this is that it is possible to hypothesise that the stochastic curves previously observed could be to some degree influenced by trends in tumour size rather than just the inherent processes underlying tumour proliferation.

In order to assess this, the time-course data previously presented was reanalysed to consider the impact of tumour size on proliferation (as measured by EdU). Reliable tumour size data are not available for the first three experiments, so these have been omitted; data comes from the latter six experiments previously discussed.

In order to visualise the impact of tumour size of proliferation, EdU uptake percentage for each tumour from each time-course is shown, subdivided by subpopulation, in the top panel of Figure 5.21. As can be seen here, there is no clear correlation of EdU labelling with tumour size, and indeed tumours of a wide variety of sizes are across the spectrum of labelling. A statistical assessment of correlation can be made via a Pearson's correlation test. A two-

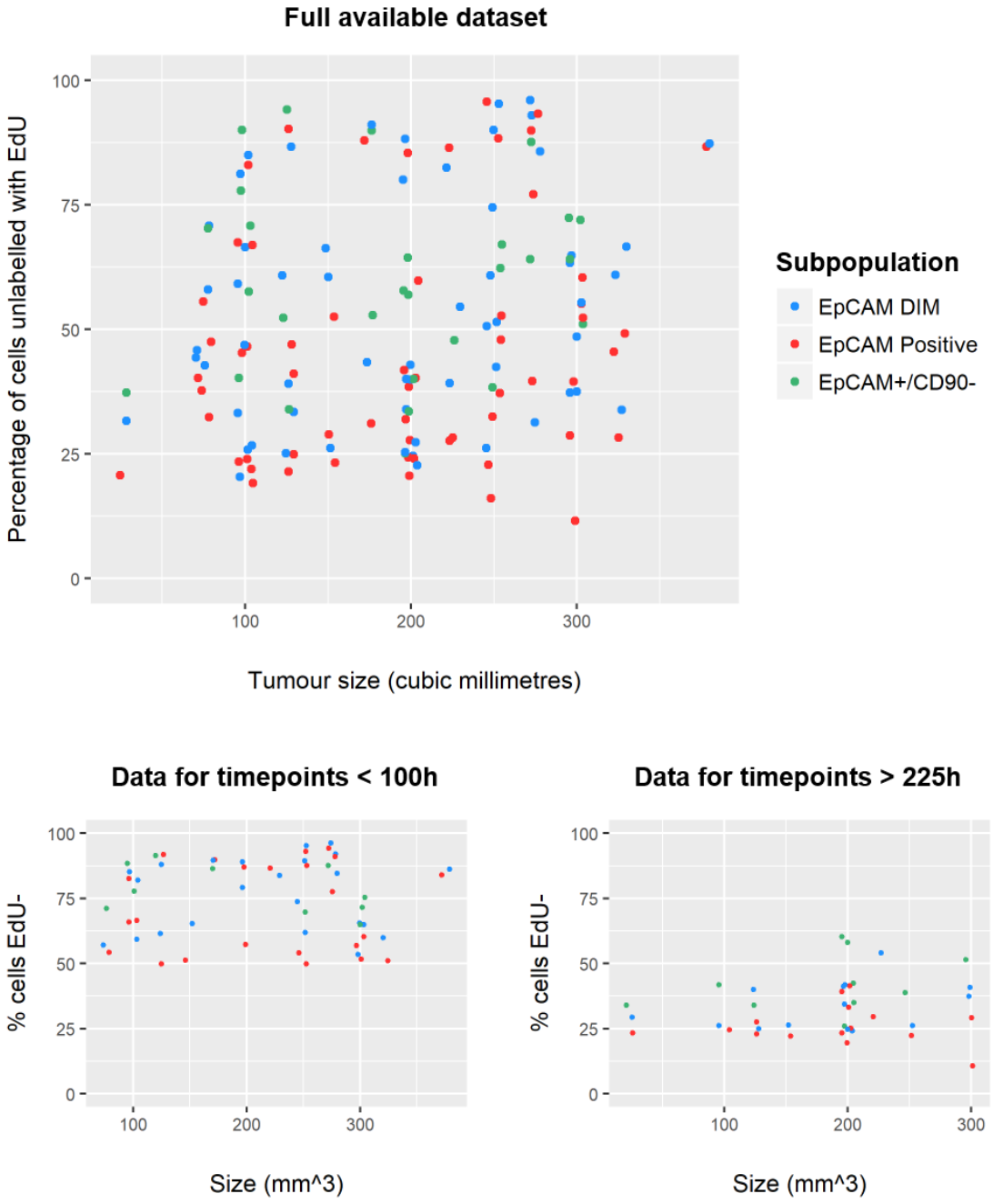


Figure 5.21: Minimal correlation between the size of experimental tumours and their EdU uptake demonstrates that size is not a major confounding factor in previous results. Top panel shows all data from previously presented time-courses for which tumour size data is available. EdU is measured as percentage of cell unlabelled to retain consistency with previous figures, but is not log transformed; size is shown as tumour volume in cubic millimetres. The top panel shows all data to show lack of a major trend overall. The bottom panels show only data for samples treated with EdU for less than 100 h and greater than 225 h respectively. The comparatively narrow horizontal spread of data shown here demonstrates a general correlation of dose time and EdU labelling, but little relation of either to tumour size. Data from $EpCAM^+$ subpopulations are shown in red, $EpCAM^{DIM}$ in blue, and (where present) $EpCAM^+/CD90^-$ in green.

sided test was used since the direction of the correlation is uncertain, and all data from each subpopulation was treated equivalently in a subpopulation-agnostic manner. This test gives an output of $r = 0.17$, where 0 indicates no correlation, and 1 a perfect correlation, indicating that that size has very little impact on EdU uptake. Curiously this test actually demonstrates statistical significance for the presence of a correlation ($p = 0.035$), since despite the correlation being very weak, the dataset is extensive enough to give low uncertainty around its presence (the p-value effectively being a measure of whether something is present, rather than if it is actually meaningful). Statistical purists might challenge the use of a Pearson's test, since it assumes normally distributed data, rather than the even distribution seen here. A Kendall correlation avoids the reliance on such assumptions, although it does so at the cost of some sensitivity ($\tau = 0.12$, $p = 0.44$). As such I retain a preference for the former in these specific circumstances.

However, such a broad attempt to assess these data is not an optimal approach to this problem. This is because it looks at data from a broad swathe of time-points and conflates them, since the number of hours a given tumour has been dosed with EdU is not indicated on this plot. In order to compare like-for-like, this variable must be controlled for. If size truly impacts EdU uptake, then differently sized tumours treated for the same amount of time should show different rates of EdU incorporation. Since in each experiment samples were measured at slightly different time-points, data has been stratified into samples treated for less than 100 h, and samples treated for more than 225 h. These data are plotted in isolation in the bottom panels of Figure 5.21.

As can be seen in both plots, there is no discernible variation in EdU uptake rate with varying size. This indicates that once the main experiment's true independent variable (hours of EdU dosing) is controlled for, the confounding variable (tumour size) shows no impact on the rate at which EdU is taken up into tumours. In contrast, the effect of this stratification by EdU dosing period can be very clearly seen, with the data corralled into the top and bottom halves of the graph respectively.

Repeating the Persons' test on these data sets once the effect of EdU dosing time is stratified for (even if not completely controlled), gives a correlation coefficient of $r = 0.0017$ ($p = 0.99$) for times less than 100 h and $r = 0.16$ ($p = 0.34$) for those over 225 h (assessment of the third fraction which falls in between these two groups, for the sake of completeness, gives a coefficient of $r = 0.13$ ($p = 0.93$).

Based on this assessment it can be reasonably concluded that the use of differently sized tumours does not have a significant impact on the proliferation values being reported within the main experiment. As differently sized tumours incorporate EdU at similar rates when exposed to EdU for the same period, it is hard to argue that larger tumours inhibit EdU uptake, and that therefore the trends shown in the main experiment are questionable. Given this conclusion when dosing time is controlled for, it is likely that that weak but statistically significant correlation seen in the combined dataset is a result of some form of covariance, although the exact origin is unclear. Equally, statistical significance in correlation does

not equate to meaningful impact. As an approximate measure of this, a subset of smaller tumours (defined as between $50\,mm^3$ and $150\,mm^3$) and a subset of larger tumours (defined as between $250\,mm^3$ and $350\,mm^3$) were compared based on their distributions of values for percentage of cells unlabelled with EdU. While the larger tumours do indeed show a slightly higher median, even subsetting and comparing extremes like this isn't enough to show a statistically significant difference at a level $p < 0.05$, when compared using a one-sided Student t-test.

In conclusion, tumour size is an important potential confounding variable, but investigation shows a very limited degree of correlation of EdU labelling rates with size. When compared with the real variable of EdU dose duration, the effect is striking. Even where a potential covariant correlation can be seen, the impact this has is very weak, and does not demonstrate a significant impact on the values for percentage labelling. Based on this tumour size can be neglected as a confounding variable.

Impact of tumour vascularisation and spatial heterogeneity on proliferation and EdU labelling

Heterogeneity in tumour vascularisation, and broader questions around spatial heterogeneity and its impact on tumour growth should be considered. However, the elimination of tumour size as a factor in EdU uptake goes a long way to tempering concerns around the former, as one of the chief considerations is that vascularisation could impact tumour growth (by increasing available resources), which could in turn impact EdU uptake. As tumour size has already been eliminated as a factor in EdU uptake, the concern that vascularisation has an impact via increased growth can be neglected, as if so, it would have been detected indirectly in the size data discussed above.

This does not however preclude a situation in which a varying level of vascularisation has some impact on growth rate by means independent of tumour size (for example through signalling factors transported via the vasculature, or through more short-range signalling between tumour cells and the murine stroma which forms the vasculature). However, if this is the case it is unclear how these largely intratumoural effects would impact the aggregated (i.e. average across a whole PDX) EdU measurements made in the experiments previously described. For this to be the case, tumours dosed with EdU for longer would need to have consistently higher or lower levels of vascularisation either by sheer statistical chance (which is extraordinarily unlikely) or via the direct influence of EdU (which will be assessed and excluded shortly. As such, while variation in vascularisation is entirely plausible, such a situation would not meaningfully impact the trends identified, or the conclusions drawn from the previous experiments.

Crude observation of PDX samples upon surgical removal from the mouse post-mortem were similar across the tumours; a solid spheroidal tumour typically with one or more blood vessels extending through the dermal layer before branching to envelop the tumour, in a fashion

reminiscent of the classic diagram of an alveolus (although much more sparsely vascularised). In some samples, aberrant vascularisation was seen, with knots of blood vessels adjacent to the tumour, wrapped around fibrous or stromal structures, and the tumour itself poorly vascularised. In these cases, tumour growth tended to be poor (which supports the assertion that good vascularisation is essential for PDX growth), and these tumours grew slowly and unpredictably, and as a result were excluded from the bulk of subsequent experiments. It is interesting to note that aberrantly vascularised tumours typically gave rise to other aberrantly vascularised tumours upon their eventual passage, and that tumours where this effect was not seen did not give rise to it in subsequent passages. This strongly suggests this effect is sample-specific, indicating that samples have varying ability to recruit vasculature (in the murine setting at least).

As such there is no need to consider poorly vascularising samples as a confounding factor as they self-select themselves out of the experimental pool, however it represents a potential limitation of the PDX methodology, in that this work inherently is a study of only the subset of HGSOC tumours which will efficiently passage upon xenotransplantation (including ability to recruit a minimum level of supporting vasculature).

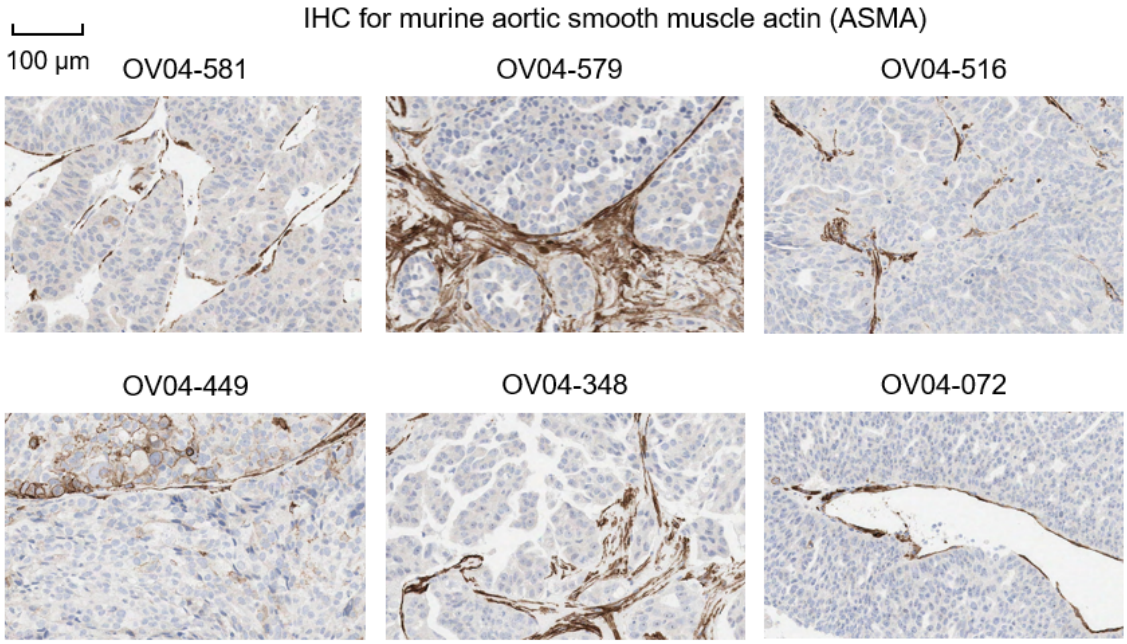


Figure 5.22: Immunohistochemistry staining for distribution of stromal cells via murine aortic smooth muscle actin (ASMA).

Samples of all the tumours used as part of the main experiments previously discussed were banked for subsequent histological assessment. Immunohistochemistry (IHC) can be used to assay vascularisation within tissue sections from these samples. IHC was conducted for

murine CD31 and murine aortic smooth muscle actin (ASMA). Both markers identified the same mix of stromal tissues, but ASMA gave a much better signal. Example ASMA staining is shown in Figure 5.22.

Among the tissues identified, some are clearly elastic or fibrous connective strands (e.g. as shown in the top middle image of Figure 5.22), while some appear consistent with vasculature (e.g. bottom right panel). However, there are also examples (e.g. top left) of what appear to be thin layers of vascular tissue, but where their distribution and layout are not consistent with vessels, but appear to trace arbitrary outlines of apparent pseudo-structures. Equally, some tumours appear to contain large luminal spaces, surrounded by thin layers consistent with vasculature, where the space appears far too large or too branching to be consistent with a functioning vessel. This suggests that in many cases, rather than vasculature, these layers of $ASMA^{+}$ may represent disorganised recruitment of murine stromal tissue to the tumour as a result of aberrant signalling from the tumour cells, or the impact of tumour cells growing in a heterotopic environment.

Given the degree of uncertainty around these various types of structures, quantitative assessment is not likely to be particularly helpful (as different things would be being measured in each case). However qualitative review of the IHC data indicates that vascularisation is sparse across all of the samples. This is consistent with crude observations during the experimental work; that cutting into the vasculature surrounding the tumour releases blood and fluid, while cutting into the hard bulk of the tumour itself does not result in significant bleeding.

To further investigate this phenomenon, the same samples were also stained by IHC for the human form of the proliferation marker Ki-67, to see if proliferation measured by this surrogate meaningfully co-localised with any of the spatial, vascular, or stromal structures described above.

One striking feature which can be observed in several tumours is exemplified in the top panels of Figure 5.23. These panels are from the same slide and captured under identical conditions, however the top-left panel is taken from near the centre of the PDX and the top-right is taken from the edge. As can be seen, the periphery stains much more homogenously than the core of the tumour, where a smaller proportion of cells appear to stain positive for Ki-67, indicating that proliferation is higher on the periphery of the tumour than it is in the core.

It is worth reviewing these images carefully, as there appears to also be a confounding effect at work. The cytoplasm of the cells at the edge of the tumour appears to stain much more darkly than the cells in the centre, making this image look much more strongly stained at first glance. This suggests that the core of the tumour may also suffer from reagent limitation in the IHC assay, making the peripheral sample look more Ki-67 positive than is actually warranted. To mitigate this, it is worth considering only the density of the most strongly stained nuclei. By this measure, the periphery appears to have around twice as

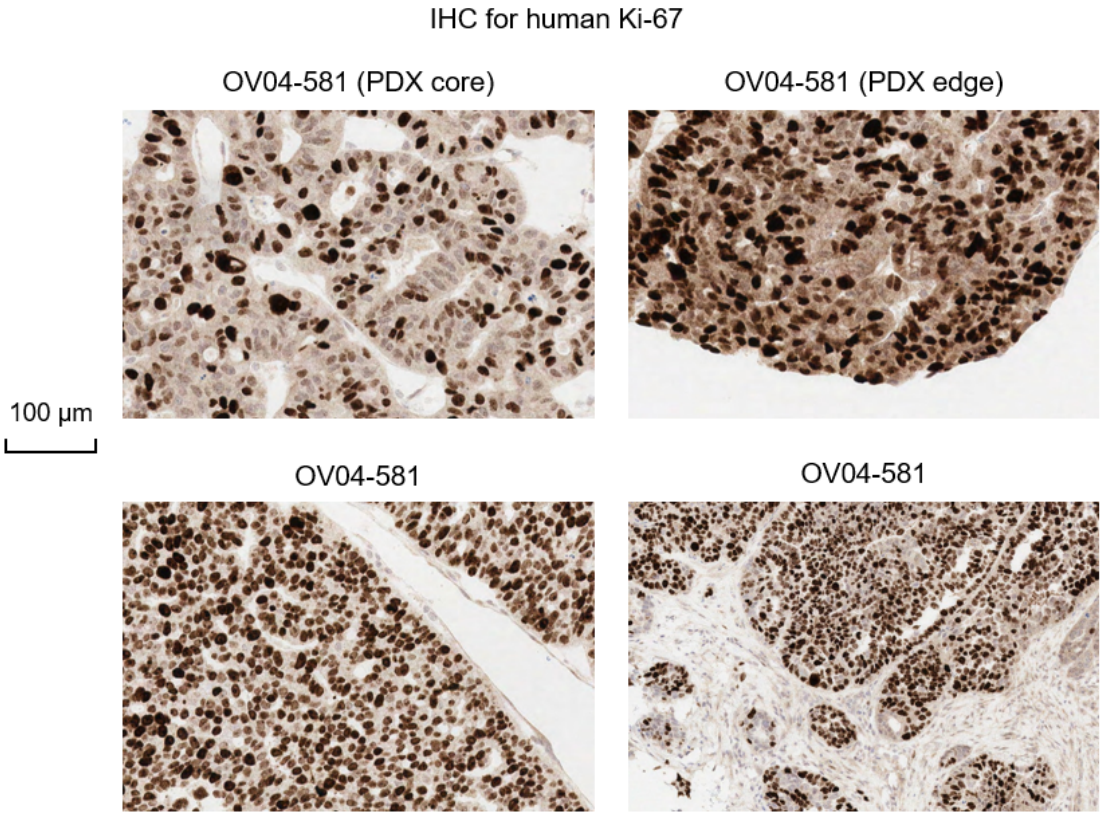


Figure 5.23: Immunohistochemistry staining of PDX samples for proliferation marker human Ki-67.

many proliferative cells as the core, suggesting that the tumour is growing much faster at the edge than in the centre.

This outcome is perhaps intuitive, as cells closer to the edge of the tumour will have easier access to resources being transported via external vasculature, as well as to any signalling factors transported in the bloodstream. Cells around the edge of the tumour may also be less crowded, and subject to anti-growth signals induced by pressure and cell-cell contacts which indicate 'crowding' (assuming that the transformed cells are still sensitive to such stimuli). This spatial effect was not seen universally, indicating that such spatial effects are subject to a range of tumour specific features.

The second question addressed by this IHC work is whether stromal and vascular features affect proliferation. Such effects were not seen in any of the 12 PDX samples assessed by IHC. The bottom panels of Figure 5.23 show examples of a pseudovascular vessel-like structure and dense elastic stroma respectively; there is no sign of proliferation rates (as measured by Ki-67 positivity) varying in the vicinity of these structures, either at the scale shown in these panels, or when looking at broad regions of tumours characterised by these features. Another 'structural' effect where no impact on proliferation is observed is the pseudo-fallopian structures exemplified by the top-left panel of Figure 5.23. In samples where these features were observed, these areas showed no difference in Ki-67 staining compared to areas of the tumour where no structures were apparent. The only impact of proliferation observed was due to proximity to the edge of the tumour as previously discussed, and pseudo-fallopian structures were observed in both peripheral and core regions with the expected proliferation differential based on this localisation alone.

One potentially significant consideration was raised by a piece of work by the laboratory of Cédric Blanpain, published after the previously described work was conducted (Pastushenko et al., 2018). This paper describes studies of EMT in germline mouse models of breast cancer, provides strong evidence for a continuum of EMT states in this cancer type, and reports that these states may be heterogeneously distributed.

Qualitative review of IHC for EpCAM, and cytokeratins 7 and 18 found no significant heterogeneity in these markers within the viable epithelial cell fraction (murine stromal cells stained only at background levels, murine fatty tissue did not stain at all, and small areas of necrotic tumour cells showed strong non-specific staining). There were only two exceptions to this, which will be discussed shortly. The lack of heterogeneity in the markers of interest in the project indicates that there does not appear to be a significant spatial component to the results previously discussed. It should be noted that this represents absence of evidence rather than evidence of absence, and the lack of a major spatial component to the features discussed in this work does not preclude the potential for significant spatial heterogeneity in other phenotypic markers. However, it does indicate that the results presented previously are not significantly confounded by the effects of any spatial heterogeneity which may exist; investigation into other types of heterogeneity (for example the range of potential EMT indicators discussed in Pastushenko *et al.*) is beyond the scope of this work.

The first of the two exceptions noted above was OV04-536, a sample excluded during validation for questionable confirmation of HGSOC (no *TP53* mutation), but accidentally included in this IHC run. This error was identified at the beginning of the analysis phase, and the sample neglected for the purposes of further assessment. However, this sample notably showed a marked heterogeneity for CK7, which was not seen in any of the HGSOC samples, and is worth noting as it acts as an inadvertent negative control, demonstrating that the CK7 IHC approach is sensitive to heterogeneity but that such heterogeneity simply isn't present in the HGSOC samples.

The second exception was found in the EpCAM IHC staining of sample OV04-449. In this sample there are 3 clusters of cells which stain considerably more strongly for EpCAM than the bulk of the epithelial tissue. The three clusters are widely dispersed across the tumour, but each stain in a similar fashion, suggesting the possibility of a common origin. Each is a contiguous patch, which implies some degree of clonality. The patches combined represent only a tiny part of the tumour (around 1-5% of total epithelial cells). Example staining for a matched area is shown in Figure 5.24.

The broader OV04-449 sample is clearly epithelial by both morphology and CK18 staining. It has also been confirmed to be HGSOC by virtue of IHC for p53 and WT-1 (Figure 3.6), *TP53* sequencing by TAm-Seq (Table 3.3), and the characteristic genomic fragmentation observed via shallow whole genome sequencing (Figure 3.10). The aberrant patches show human epithelial staining patterns, indicating they clearly do not represent murine contamination, however these patterns are inconsistent with the intensities seen in surrounding tissue. This is curious, and suggests either an intrinsic heterogeneity not seen in any other samples, or potential contamination. If this does represent contamination, this could be either technical, in which OV04-449 has been contaminated with another HGSOC sample at low level during its time in the laboratory, or a more aetiological process in which two distinct tumours have arisen in the same patient (for example, divergent clones from the same precancerous transformation) and a mix has been obtained during sample collection. The existence of this latter phenomenon was previously observed in OV04-404 in which HGSOC was mixed in with a predominant non-HGSOC tumour which resulted in this sample being excluded from further use in this study. WT-1 staining was performed on this sample, and is shown in the bottom-left of Figure 5.24. The consistent positive staining for this marker confirms that both areas of cells are HGSOC tumour.

These patches are also interesting in that something that matches their parameters was noted during the original assessments of EdU staining by flow cytometry. In OV04-449, and no other samples, in some cases a small cluster of events was noted appearing to overlap the top of the main $EpCAM^+$ subpopulation. This cluster was very small and very poorly differentiated from the broader $EpCAM^+$ fraction, and as such it was challenging to be confident it was truly present, and impossible to accurately separate the two subpopulations (as the centre of density of this unusual subfraction was well within the upper bound of the $EpCAM^+$ subpopulation). While not possible to fully distinguish, EdU uptake for this tiny

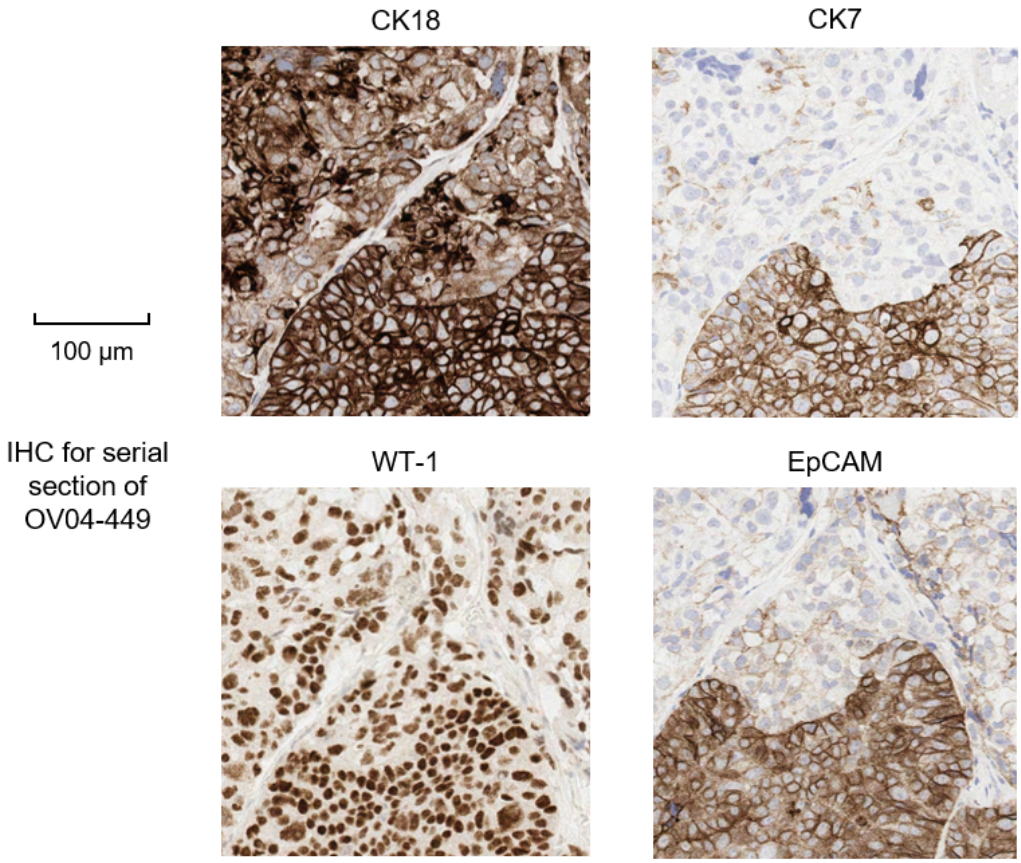


Figure 5.24: Immunohistochemistry staining for EpCAM, cytokeratin 7 (CK7), cytokeratin 18 (CK18), and WT-1 for sample OV04-449 highlights three small clusters of cells (a single cluster is pictured) which appear phenotypically distinct from the rest of the tumour, a phenomenon not seen in any of the other samples. The four panels shown here show the same area of the tumour for each of the four markers. The darker area for CK7 and EpCAM (and to a lesser extent CK18) is the rare subpopulation (around 1-5% of total epithelial cells) while the bulk of the tumour resembles the top/top-left of the images shown. Both areas are positive for WT-1 indicating all cells are human tumour.

subfraction appeared to be lower than for both the $EpCAM^+$ and $EpCAM^{DIM}$ subpopulations. Due to these complications, and that the subpopulation was too small to effectively analyse, these events were neglected in the analysis. However, in hindsight it is interesting to note that their characteristics are somewhat in line with what was observed from the $EpCAM^+/CD90^-$ subpopulation in the 3-subpopulation samples (although noticeably higher EpCAM was not observed in these samples). That they are contaminants remains a possibility, but it is also possible that they represent a minor subpopulation along the lines of $EpCAM^+/CD90^-$. This possibility is extremely interesting in that, if true, the apparent segregated clonal distribution would support the idea that this phenotypic subpopulation is self-supporting rather than being driven by stem-cell mechanics. However, without excluding contamination from another HGSOC sample (something that is extremely challenging for groups of cells this small), no definitive conclusions can be drawn from this observation. In either case, this unknown subpopulation was excluded from consideration in the work described previously in this chapter, and therefore cannot have had a confounding effect. Even if included, these cells represent a sufficiently small fraction of the $EpCAM^+$ subpopulation that it would not meaningfully change the overall trends observed even if this subpopultion were incorporated into the $EpCAM^+$ fraction. However, the observation noted above indicates the possibility of using alternative non-EpCAM markers to identify and demonstrate clonal proliferation within phenotypic complexity. This is a potential avenue for future investigation.

However, it is worth noting that the spatial heterogeneity discussed in Pastushenko *et al.* was observed in germline-modified mouse models of disease; the disruption of the human tumour by surgical removal (or via the tumour metastasis which is presumed to create ascites) followed by transplantation into the PDX host may have disrupted any existing spatial heterogeneity which may not have been recapitulated. Alternatively, since the PDXs are created from a very small piece of tissue, only a single part of a spatially heterogeneous tumour may have been transplanted. As such it is reasonable to conclude that while the previous work has not been confounded by any significant spatial or vascular effects, but that limitations of the model system mean that the full possibilities of spatial or vascular heterogeneity possible in the original patient tumour have not been fully recapitulated, and that this PDX model may not be ideal for consideration of this element of possible complexity.

In conclusion, qualitative review of vascularisation and spatial considerations with regard to the EdU labelling work discussed previously shows little meaningful correlation, and as such there does not appear to be a major spatial component. The only noticeable feature was the proliferation difference between the periphery and core of PDX tumours. As every tumour assessed contains both, the aggregated average can be presumed to be representative (as the EdU experiments assess proliferation at a whole-tumour level). The only factors which could influence this would be shape and size, as these elements could influence the comparative proportions of the two areas. Shape can be discarded as the consistently growing samples used in the EdU work were all broadly spheroidal, while size has already been assessed and

neglected as a confounding variable for proliferation. This indicates that the previous work is not confounded by these concerns.

EdU does not selectively inhibit proliferation and EdU labelling

All previously shown long-term EdU experiments are subject to the assumption that EdU itself is not significantly influencing the proliferation rate of the two populations. Since EdU involves the incorporation of a nucleoside analogue into DNA (i.e. a similar but not strictly canonical structure being introduced to an enzyme driven biological process), this assumption is open to question, since once marked with EdU a cell's cycling time could be slowed or its ability to proliferate could even be inhibited due to interruption of the DNA replication machinery. Since short-term EdU studies measure EdU incorporation only and measure only whether cells are in S-phase at the time of dosing (not whether they subsequently complete it normally), they should be less likely to be subject to any effect, but this possibility cannot be strictly excluded.

As nucleoside uptake experiments use the presence of the analogue to detect proliferation, a true negative control is not possible. However, the effects of EdU on division rates may be inferred by applying varying concentrations of EdU to different tumours, experimentally analysed under identical conditions, and using the results to titrate a rate of labelling against concentration. To do this, a cohort of three mice was set up (passaged from the same tumour) and given EdU-treated drinking water for 5 days at concentrations of $0.3\,\mathrm{mg\,ml^{-1}}$, $1\,\mathrm{mg\,ml^{-1}}$, and $3\,\mathrm{mg\,ml^{-1}}$ respectively. On day 5 the mice were killed and tumours harvested and assayed for EdU. The data from this experiment are shown in Figure 5.25. This time-point was chosen as a compromise between minimising variation and maximising signal. If too early a time-point is used, there is a risk that variation in time at which mice first drink (initiating marking) outweighs the signal. If too late a time-point is used, the signal itself becomes much weaker because most of the cells are already marked, and a weaker signal is less pronounced against background noise.

As can be seen, the rate of incorporation appears to be linear and concentration-dependent. However, the drop in the proportion of cells marked is similar in both subpopulations, versus the extrapolated value for zero concentration. For the $1\,\mathrm{mg\,ml^{-1}}$ used for all the EdU time-course experiments, these data suggest proliferation is slowed by 10.4% for the $\mathrm{EpCAM^{+}}$ subpopulation, and 9.5% for the $\mathrm{EpCAM^{DIM}}$ subpopulation. As such, this experiment supports the validity of the previous conclusions regarding the relative rates of proliferation in the various subpopulations. However, it suggests that in order to accurate estimate absolute DTs, all percentages should be divided by a factor of 1.1 in order to extrapolate the true division rate minus the effects of EdU. Thus a time-point reporting 55% of cells remaining unmarked, should be read as 50% unmarked (45% of cells in the assay are marked, but EdU slows proliferation by 10%, so minus the effects of EdU this is actually 50%).

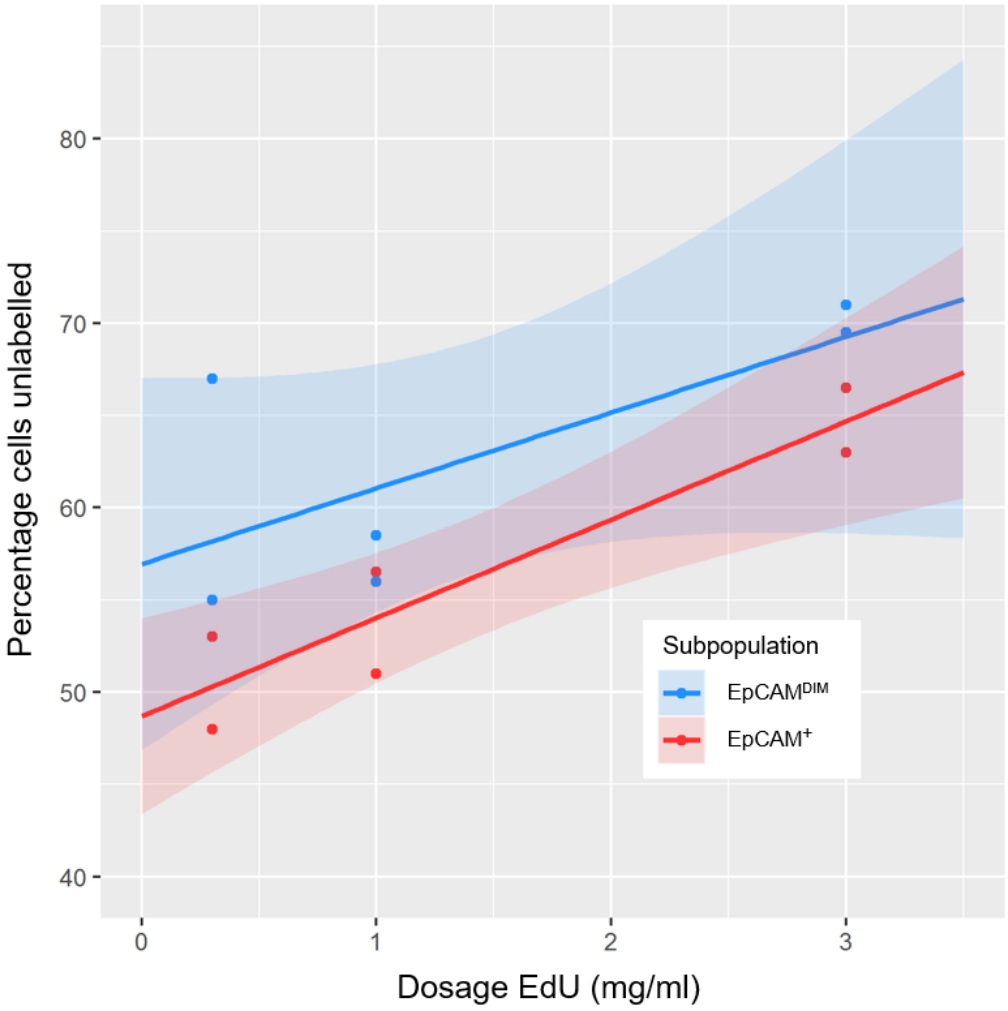


Figure 5.25: A titration of different concentrations of EdU, done in parallel across the same experiment. EdU slows proliferation rate (here measured by EdU incorporation), but does so in a linear concentration dependent manner. Reduction in proliferation (measured in relative increase in percentage of cells remaining unlabelled) is similar for both $EpCAM^{+}$ (red), and $EpCAM^{DIM}$ (blue). This suggests EdU does indeed inhibit proliferation but in an unbiased fashion. These data support the validity of the primary tracing experiment.

5.4 Discussion

The objective of this chapter of work was to determine the *in situ* proliferation patterns associated with the previously identified subpopulations within the HGSOC PDX tumours, and where appropriate to estimate average proliferation rates for these.

The data presented within this chapter show that uptake of EdU during long-term division assays occurs, in all subpopulations, in a manner consistent with stochastic loss from the unlabelled pool.

The median division times estimated from short-term EdU labelling and long-term EdU labelling experiments are broadly similar, and show very similar relative rates of division between the subpopulations. This is instructive because it indicates that the rate of proliferation within that subpopulation (as assayed by the short-term EdU trace) is similar to the rate of new cells accumulating in that subpopulation (assayed by the long-term EdU trace), which is indicative of subpopulations being self-supporting. It cannot be conclusively proved from these experiments that the subpopulations are not exchanging cells, however if this is the case, this must be either occurring at too low a level to be detected in these experiments, or the process must be tightly controlled in some way.

These experiments have shown that in most cases there are two apparent subpopulations, an $EpCAM^{DIM}$ and an $EpCAM^{+}$, and that the former appears to cycle up to twice as slowly as the $EpCAM^{+}$. There are also indications from at least one sample, that in some cases this distinction can be diminished to the point where the two populations converge. Further research would be needed to understand in more detail what parameters are able to result in this unusual result.

In addition, these experiments have allowed the estimation of median division times for each population. These calculations can be performed based on both short-term and long-term EdU assays. The values derived by each are similar, despite the independence of the two methods, suggesting a relatively high degree of confidence that they are broadly accurate.

In some cases, these experiments show three distinct subpopulations, with an $EpCAM^{+}/CD90^{-}$ resolving from the main $EpCAM^{+}$ subpopulation which is $CD90^{+}$. It is not clear if this $CD90^{-}$ population is absent in other cases, or if it is just less distinct from the main $EpCAM^{+}$ subpopulation, and cannot be resolved. In cases where CD90 expression is low in the population overall, this explanation seems very likely. However, in some cases where CD90 is highly expressed it seems curious that such a subpopulation does not resolve if it does indeed exist.

The proliferation data in these cases appears much less clear. While the $EpCAM^{DIM}$ subpopulation in all cases does proliferate slower than the $EpCAM^{+}$, the difference is too small to be significant. However, in these cases the $CD90^{-}$ subpopulation does appear to proliferate significantly more slowly than the $EpCAM^{+}$ bulk. It is not clear what the explanation for

these results is. Possibly this subpopulation represents a fraction of the bulk $EpCAM^{+}$ which have altered their cell state, downregulating some markers (including CD90) and reducing proliferation. However, if so, it is unclear why in each case the $EpCAM^{DIM}$ population has also lost its distinctive slow-cycling characteristics at the same time. An alternative hypothesis would be that these are in some way related to the $EpCAM^{DIM}$ subpopulation, and that their emergence is in some way covariant with the loss of distinction between the $EpCAM^{+}$ and $EpCAM^{DIM}$ subpopulations. As things stand there is too much phenotypic complexity to determine any additional information of value by studying proliferation characteristics in this way. Further interrogation of this phenomenon is likely to require further characterisation of the subpopulations at the phenotypic level. Mechanistic information from such studies can be used to further understand the relationships between the populations. Once clearer hypotheses have been generated, these can be subsequently tested using further proliferation assays. Initial work to this end is discussed in the next chapter.

It is theoretically possible that a tiny proportion of cancer-stem-like-cells seeded within a much larger stochastically dividing $EpCAM^{DIM}$ population could produce similar results. However, for this to be the case, these cells would need to not contribute significantly to the process of maintaining / growing the tumour, raising questions of their relevance, and the rationale for their existence. Further investigation of these phenotypic subpopulations depends on a better understanding of possible exchange between these subpopulations, as well as the underlying mechanistic character of these cells. This will be discussed in the next chapter.

Chapter 6

Characterising cell fate in the context of phenotypic subpopulations

6.1 Introduction

The work discussed in the previous chapter looked at average rates of proliferation within different phenotypic subpopulations, both over the short-term and the long-term. These studies also determined that the average rates of short-term proliferation in each subpopulation matched the average rates of long-term EdU accumulation in each subpopulation, and concluded from this that each subpopulation was likely to be self-supporting.

However, all this work strictly shows is that, on average, there is no *net* movement of cells between subpopulations. There could be anywhere from extensive (but equivalent) interconversion, to none at all. The likelihood of there being extensive interconversion would seem to be fairly low given that in order for equivalence to be maintained, some kind of regulatory process would be required (or else the long-term EdU marking rates of the two populations would be expected to equilibrate). However, low-level interconversion or no interconversion between $EpCAM^{+}$ and $EpCAM^{DIM}$ states could easily happen without the effects being large enough to show up as significant differences between the short-term and long-term EdU datasets.
It is already clear from Siru's work that there is apparent interchange of $EpCAM^{DIM}$ cells into $EpCAM^{+}$ cells during tumour initiation, given that a pure $EpCAM^{DIM}$ population appears to be able to give rise to a phenotypically diverse tumour. However, this occurs during the extreme stress of xenotransplantation, and may represent an aberrant 'wound-healing' type response. What is not clear is whether this type of event occurs normally in

established tumours or PDX xenografts, and if so at what rate.

Within the larger translational context of the project this otherwise minor detail is extremely significant. The $EpCAM^{DIM}$ subpopulation shows all the properties associated with a minimal residual disease subpopulation (slow cycling, chemoresistant, loss of epithelial character); it appears likely that it is this population which is surviving conventional carboplatin therapy and repopulating heterogeneous disease through potential wound-healing responses. However, additional information about how able cells of one phenotype are to give rise to progeny of another (or how able existing cells are to interconvert) under normal conditions opens up new approaches to the development of future treatments.
If the $EpCAM^{DIM}$ subpopulation is self-populating, it may be possible to target drugs against this minimal residual disease subpopulation prior to chemotherapy. By killing off the resistant cells first, carboplatin treatment could subsequently be used to eliminate the bulk disease. However, if it is possible for cells to quickly interconvert, such an approach would be undermined since the bulk disease would be able to regenerate the minimal residual disease subpopulation between treatments. In such a case a better approach might be to take advantage of this very plasticity, targeting a drug to disrupt regulation of this process and promote conversion of cells from $EpCAM^{DIM}$ to $EpCAM^{+}$ to allow for elimination of all cells using conventional carboplatin treatment.

At present, all the work described previously has looked at cellular subpopulations separated by differential expression of one or two key markers using flow cytometry, along with a number of biological characteristics (proliferation rate, chemosensitivity). For any subsequent drug development work to be targeted at these populations, a much more detailed understanding of the actual molecular biology within these cells will be needed. Such information also allows for further elucidation of the links between the subpopulations in question.
Two working hypotheses, generated during the previous chapter of work, regarding the identity of the $EpCAM^{DIM}$ population (aside from that it is slow-cycling) is that it might represent a dedifferentiated subset of the bulk of epithelial cells, or a subset that has undergone epithelial-to-mesenchymal transition (EMT). Both of these possibilities would account for the reduction in EpCAM expression. Further investigation of the expression patterns of these populations would allow for testing of these hypotheses, and also allow light to be shed on the nature of the $EpCAM^{+}/CD90^{-}$ subpopulation seen in some samples, the nature of which is at present unclear.

The aims of the following work are therefore three-fold:

- To determine if cells are able to produce progeny of a different subpopulation, or otherwise interconvert
- To estimate a rate for such a process (if it is present)
- To determine the expression patterns and understand the biology of the subpopulations previously indentified

6.2 Methods

6.2.1 Lineage tracing

One method for determining the fate potential of cells is lineage tracing. This involves a small number of cells being marked at the gene level, and the tissue/tumour being allowed to grow. Marked cells which proliferate will give rise to clonal patches of cells all containing the mark. By studying these patches for the presence of multiple lineages, the presence of interconversion or generation of progeny with different lineages can be established. It is an assumption of the method that marking is sparse enough that multiple clonal patches will not converge, as this will lead to aberrant results. Typically this method uses mouse models into which markers have been introduced at germline level, but which require induction to be expressed. In order to initiate marking, the appropriate inducer is added at a carefully titrated concentration in order to sparsely mark cells. If this marker is not tightly regulated, aberrant non-specific induction can occur too early leading to marking of vast swathes of cells which is of no use experimentally.

In the context of this work, germline marking is not possible since the PDX material is of human origin. In order to mark cells for lineage tracing in this case, a PDX tumour must be processed down to single cells, the cells virally transduced with an inducible marker which can be activated by a titrated inducer in the usual fashion, and the transduced cells re-injected in order to form a new PDX tumour. Viral transduction of a marker introduces a number of additional challenges into this experiment. Firstly, a suitable viral construct is needed. This construct must express a fluorescent marker which is only activated in response to induction. For practicality purposes a second constitutive marker is generally included so that the rate of transduction can be assessed without induction. Secondly, this construct must show minimal spontaneous induction, or by the time the final PDX is induced, many cells will already have aberrantly induced and it will not be possible to distinguish induced clonal patches amid swathes of autoinduction. Thirdly, HGSOC cells survive poorly *in vitro* and so a transduction methodology must be optimised in order that minimises the time the cells spend *in vitro* while ensuring a sufficient degree of cell marking. Fourthly a sufficient degree of transduction must be achieved such that a good proportion of the final PDX tumour cells are marked. This is contingent not only on the efficiency of the original transduction, but also on the DNA insert not being deleterious to survival of host cells over the long-term.

Once a PDX has been marked, and a suitable concentration of inducer determined by titration, the tumour can be induced, allowed to grow for a suitable period of weeks, then formalin fixed and imaged for fluorescence. Each patch of fluorescence corresponding to the virally induced mark can then be assessed. Information about proliferation rate can be extrapolated from the sizes of the various cell patches, since these reveal total number of cell divisions since the original single induced cell. Samples can also be co-stained with antibod-

ies to antigens of interest (in this case EpCAM being the obvious candidate). By studying the proportion of patches containing both $EpCAM^{+}$ and $EpCAM^{DIM}$ cells (and how this changes depending on the size of the patch) rates of interconversion can be extrapolated.

6.2.2 Symmetry assays

Lineage tracing cannot distinguish between a cell dividing to give rise to a cell of its own type and one of another type, and a cell dividing to give rise to two cells of its own type, one of which subsequently converts from one phenotype to another. In order to assay this, symmetry can be assessed immediately after division. PDX cells can be marked with BrdU by intraperitoneal injection (using methods as per the short-term EdU trace), followed by formalin fixation 6 h later, such that a significant minority of marked cells have divided. These cells can be identified by their proximity and identical degree of BrdU concentration (due to the semi-conservative nature of DNA replication). By co-immunostaining for BrdU and EpCAM, a large number of individual divisions can be assessed to determine if they are symmetric or asymmetric, and a rate of asymmetry determined.

6.2.3 Single-cell mRNA Sequencing

From the data presented by Siru, it appears clear that the $EpCAM^{DIM}$ population is both tumour initiating and at least preferentially enriched by chemotherapy. This suggests these cells are responsible for tumour relapse after chemotherapy. If this phenotypic subpopulation could be effectively repressed, then disease recurrence should be inhibited. Given the data from the previous chapter, it seems clear that the two phenotypic subpopulations are at least largely self-supporting. Thus if the $EpCAM^{DIM}$ population could be targeted and eliminated or depleted, then repopulation of this fraction by the $EpCAM^{+}$ subpopulation should be limited. This makes the $EpCAM^{DIM}$ subpopulation an attractive target for research into new therapeutics.

If the constitutive lineage-tracing data from Janzen *et al.* (Janzen et al., 2015) is indeed identifying the same population as our laboratory (based on the very similar growth characteristics observed) then this may further support this hypothesis. The authors noted that constitutively labelled CA-125^{+} cells (presumably the $EpCAM^{+}$ subpopulation) gave rise to only their own type of cells after xenograft passage. This is consistent with the idea that the conditions/samples they used were more tolerant and allowed some $EpCAM^{+}$ cells to successfully engraft. It also suggests that the ability of their 'bulk' chemosensitive population to undergo some form of transdifferentiation (or dedifferentiation) into the chemoresistant phenotype is extremely limited.

Since the $EpCAM^{DIM}$ population appears to be an attractive target for future study with a view to therapeutic targeting, it would be valuable to understand more of the underlying

biology in order to guide future studies. EpCAM itself is simply a marker indicating the epithelial nature of a cell. It is highly improbable that loss of this marker is in any way directly causing of any of the phenotypic characteristics observed. Instead it is presumably a covariant indicator that the cell is undergoing some kind of state-change away from a highly epithelial character. Working hypotheses are that it may either be undergoing epithelial-to-mesenchymal transition, or alternatively some more general dedifferentiation process into a cell which is less strongly defined, and which may regain broader potency not normally available.

In order to understand these $EpCAM^{DIM}$ cells at a molecular level, rather than simply their biological properties (in terms of their proliferation and chemoresistance), single-cell mRNA sequencing (scRNA-Seq) was performed. The potential of this method was underscored in a recent publication by Winterhoff *et al.* who used single-cell RNA sequencing to assess the transcriptome of 66 cells from a single ovarian tumour (Winterhoff et al., 2017). The authors noted significant heterogeneity within this population, identifying epithelial-like and non-epithelial-like subpopulations. Of particular interest, in support of this conclusion they present data showing differential EpCAM expression across the two subpopulations. This publication reports the EpCAM-low subpopulation to be stromal-like in nature, and express a number of markers linked to stemness. However, one major concern with this work is that stroma was not depleted prior selection of cells for sequencing. A fairly extensive depletion of immune cells was conducted, but discrimination between tumour and contaminating stromal cells was made post-hoc on the basis of the expression data for a small number of key genes linked to HGSOC. Some of these (such as PAX8) should be reliable tumour markers, however others (such as CA-125) have been implicated in differential expression in EMT-like cells (Comamala et al., 2011) (Giannakouros et al., 2014). As such, in the absence of a conclusive method for removing stromal cells, the transcriptional differences between bulk tumour and EMT-tumour risk being lost amongst the conflating expression patterns arising from contaminating stromal cells. Thus EMT cells may be wrongly eliminated from the analysis on the grounds they downregulate markers like CA-125 and appear less like typically HGSOC. These concerns aside, this group's assessment of the prevalence of various reported CSC markers shows erratic expression of these with no particularly discernable overlap with. This would be consistent with my hypothesis that these markers are loosely associated with varying TIC capacity, and are not true CSC markers.

In addition, the number of cells used in this experiment was relatively small, limiting its power and capacity to discriminate meaningful trends. Given the novelty and cost of these techniques, this group's use of a small number of cells is entirely understandable; however this means there are untapped opportunities in similar work at a larger scale.

The development of xenograft models for this project represents an opportunity to take advantage of the natural replacement of human stroma by mouse cells. Since transcripts of mouse cells will map to a different genome, contaminating stroma cells can be definitively discarded.

Eight PDX tumours were dissociated into single cell solution, stained for EpCAM and CD90 (as well as viability using DAPI), sorted by flow cytometry, and subject to single-cell reverse-transcription polymerase chain reaction (RT-PCR) and library preparation. In order to minimise the time involved (and thus degradation of the mRNA), xenografts were processed in four batches of two.

$EpCAM^+$, $EpCAM^{DIM}$ and unsorted populations (as well as the $CD90^-$ where present) were separately prepared and in each case approximately 4500 cells resuspended in 33 µl of PBS. Downstream library preparation and sequencing was then performed by the CRUK Cambridge Institute Genomics Core. Single-cell library preparation was performed on each subpopulation (as well as the unsorted sample) using a Chromium Single Cell 3′ Kit (10x Genomics) according to the manufacturer's recommendations (Revision A). Resulting cDNA concentration was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies) according to the manufacturer's recommendations. The libraries were then pooled, using cDNA concentration from the Bioanalyser as a crude estimate of cell number for the purposes of normalisation. A single pilot lane of sequencing was then performed using a HiSeq 4000 (Illumina) in order to assess the quality of the samples, and inform normalisation for the main round of sequencing.

Library preparation

The Chromium Single Cell 3′ Kit uses microfluidics to perform single-cell mRNAseq at the single-cell level. An aliquot of cells is added to the chip along with a solution of labelling beads from the manufacturer, and oil for the microfluidics process. Each labelling bead is covered in primers for reverse transcription. Each primer includes two barcode sequences. The first is unique to the bead (i.e. all primers on that bead contain the same sequence), and so allows a cDNA sequence to be traced to its bead of origin. The second is a 10 bp randomer and thus unique to a particular primer (i.e. all primers on a given bead have a unique barcode). The microfluidics system pipes a thin stream of these labelling beads (in aqueous solution) into a larger oil stream within the reaction chip. The aqueous fluid hits the oil, and unable to mix, it forms aqueous 'bubbles' within the oil. The microfluidics and beads are specially designed such that they enter the oil at a precise rate, with a single labelling bead entering into each aqueous bubble. Just before the stream of beads hits the oil, the cell solution is piped in to join the labelling beads; the cells too are an aqueous solution, which also contains all the enzyme and reagents for downstream steps. The cell solution is very dilute, such that ∼95% of the resulting bubbles contain only a bead and no cell, while the remainder should contain one of each. Each of these bubbles, referred to as a partition, is where the reverse transcription will occur, and is identified by the barcode unique to its labelling bead. An underlying assumption of the downstream process is that each partition contains a single cell; thus the subsequent mRNAseq is not strictly 'single cell mRNA seq' but 'single partition mRNA seq'. It is for this reason that ∼95% of the partitions are left empty: if two cells enter the same partition, they will appear as a single

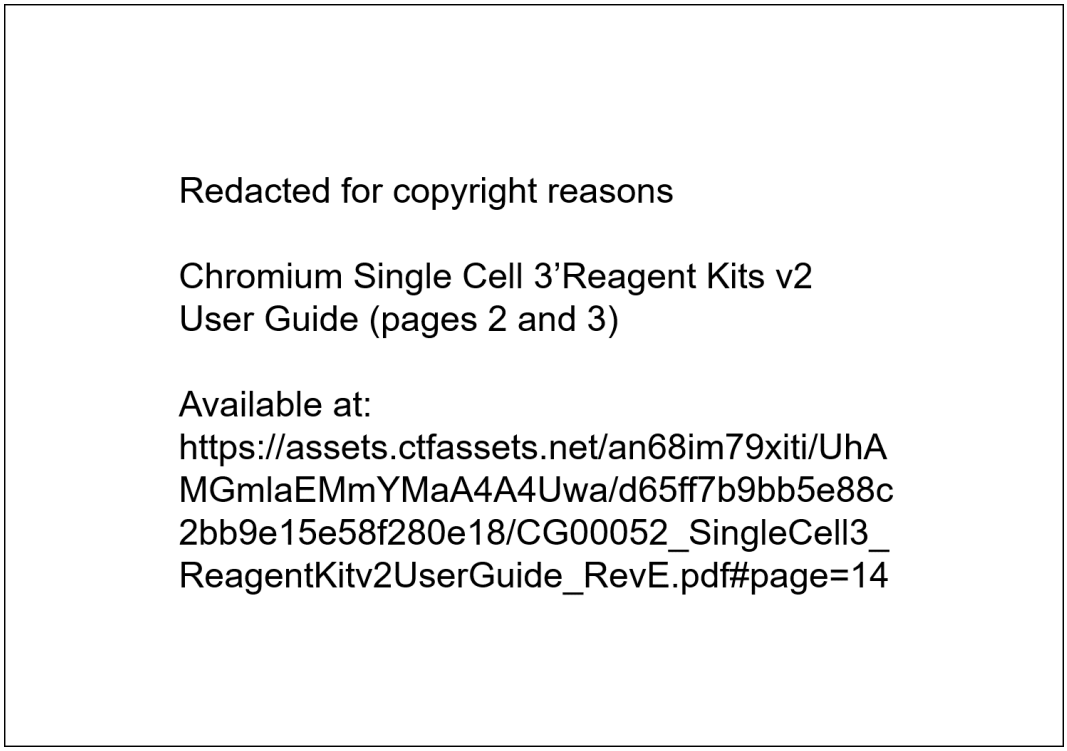


Figure 6.1: Methodology of the 10x scRNAseq library preparation and squencing

cell in the sequencing. By having an excess of partitions versus cells, the chance of two cells in the same partition (a 'multiplet') is minimised. Once this partitioning process is complete, the remaining steps follow a typical bulk mRNAseq methodology, but with the reactions occurring in isolation within their partitions.

In this case, a total of 4500 cells per sample were input. Based on manufacturer guidelines and estimated efficiency, this should yield 2500 cells for sequencing (with an estimated multiplet rate of 1.9%). After partitioning, a reverse transcription step is used to generate cDNA. A clean-up reagent then breaks up the emulsion and separates the oil. The remaining aqueous phase then undergoes magnetic purifiation (using silane beads) and ethanol washing to remove the enzyme and other transcription reagents. The cDNA is amplified by PCR (in this case 12 cycles) and quantified on the Bioanalyser (selecting only fragments between 200 bp and 9000 bp in length). Next, size selection (using a purification reagent and magnetic separator) is used to exclude very long (>9000 bp) nucleoside fragments. The purified sequences are incubated with adaptors and ligase to add an adaptor element to the cDNA, and size selected again to remove short fragments (unattached primers and adaptors). Another round of PCR (12-14 cycles), with primers containing the sample index (unique to each of the 29 biological samples used) and the P7 element is used to finish generation of the library. A final size selection step is used to purify the library, and quality is confirmed using the Bioanalyser.

Sequencing

Sequencing was performed using 26,8,0,98 paired-end sequencing on an Illumina HiSeq 4000 system. This system performs standard Illumina (Solexa) sequencing using bridge-amplification within a patterned flow-cell.

The prepared library sequences are denatured and added to the flow-cell in the presence of the reagents required for bridge-amplification (a PCR-like process). The 'patterning' on the cell consists of billions of nanoscale wells etched into the surface, which are seeded with the same adaptors (P5 and P7) as the library. The ends of the library sequences are able to bind to the adaptors in the wells, allowing extension of the bound adaptor into a full-length strand. The library sequence can then unbind from the new sequence and drift away, leaving the newly extended sequence free (but still anchored in its well by its adaptor). While anchored, this strand is still able to freely 'wave' around, and its other end is able to loop over and bind to the opposite type of adaptor bound to the surface of the same well. Where this free end contacts an adaptor, extension can take place back across the 'bridge' which has now been formed. When these strands separate, there will be two inverse anchored strands which can subsequently repeat this process. This nanowell will rapidly duplicate copies of these strands until it is full; this steric exclusion prevents further free molecules from binding and ensures each well contains only a single sequence-antisequence pair. Once complete, conditions are altered to ensure separated strands, and R1a primers are added. A mix of labelled chemically-modified terminator-nucleosides are then added along with enzyme. Each read will extend by a single base, with the appropriate nucleoside annealing and being bound by the primer. The cleavage of the nucleoside which occurs upon binding causes it to fluoresce briefly. The lack of a 3′ hydroxyl group prevents further extension. This leads to a single base extension in strands across a well. Each of these produces a flash of a colour specific to the particular base, and this causes the well to light up a particular colour during this step; this is recorded by the instrument's camera, and is translated into a letter of code. Conditions are changed to allow hydroxylation of the 3′ end, then more terminator nucleosides are added to allow another base extension. This process is repeated 26 times to capture the length of the partition barcode (16 bp) and the randomer barcode (10 bp). The Read 1b primer is then extended 8 times to read the 8 bp sample barcode which was added to allow multiplexing. The sequence from these primers allows each well to be assigned to a specific sample, cell, and unique reverse transcript. The Read 2 primer is then used to extend 98 bp in the opposite direction, to read across the cDNA insert, in order to sequence the transcript itself and determine the gene the original mRNA was encoding.

Validation

Pieces of the specific xenografts dissociated for use in each scRNAseq experiment were frozen, and subject to TAm-Seq in order to confirm that these later passage tumours correctly

matched the original patient samples. This replicated the original TAm-Seq validation process conducted when the PDXs were originally established and was to confirm that there had been no outgrowth of unrepresentative material (e.g. contaminating non-HGSOC tumour).

6.3 Results

6.3.1 Lineage tracing

Lineage tracing of PDX material first requires the ability to successfully introduce a non-transient inducible construct in order to allow permanent marking of cells, a small subset of which can then be induced to initiate sparse clonal marks. While other options were briefly investigated, the bulk of this work used lentiviral vectors due to both their proven efficiency in Siru's previous work, and that they are reported to be the most efficient way of transducing ovarian cells (Indraccolo et al., 2002). Generating a construct *de novo* is a time-consuming and involved process, so suitable constructs already available within the CRUK Cambridge Institute were investigated. The lentiviral construct LV-indLS2 (see Figure 6.2) was created by Scott Lyons for similar work (Rodriguez et al., 2014).

This construct uses a CAGGS promoter to drive expression of a Cre variant (CreERT2) which only becomes active in the presence of 4-hydroxytamoxifen (a metabolite of tamoxifen). In the absence of the inducer it expresses only inactive CreERT2, and no fluorophore. Once induced by addition of 4-hydroxytamoxifen, the now active CreERT2 cleaves the sequence at the two variant loxP sites (lox 66 and lox 71) excising the CAGGS promoter. Due to the relative strengths of these different sites, the CAGGS promoter will preferentially re-integrate in the opposite orientation. In this new orientation the CAGGS promoter drives transcription across a multicistronic vector containing a combined luciferase-mStrawberry protein (joined by a self-cleaving E2A sequence), causing the induced construct to fluoresce.

This construct was used to generate virus in 293T cells using 3rdgeneration lentiviral packaging (pMDLg/PRE, pRSV-Rev, pVSV-G) and the transfection reagent lipofectamine. Initial experiments showed very low levels of LV-indLS2, such that it was unclear if the signal was in fact real. Expression of a simple constitutive construct expressing zsGreen from an EF1α promoter (run in parallel as a control) was significant, but well below what was expected. The poor positivity for the control virus suggested a methodological issue, and so optimisation and troubleshooting was performed to boost the efficiency of the generation and subsequent transduction of virus. These assays were performed using the zsGreen control construct. These are summarised in Figure 6.3.

A 3rdgeneration lentiviral system as previously used is less efficient than 2ndgeneration systems (the latter using two rather than three packaging constructs in addition to the payload construct). Reducing the number of plasmids that must be taken up by the virus generating

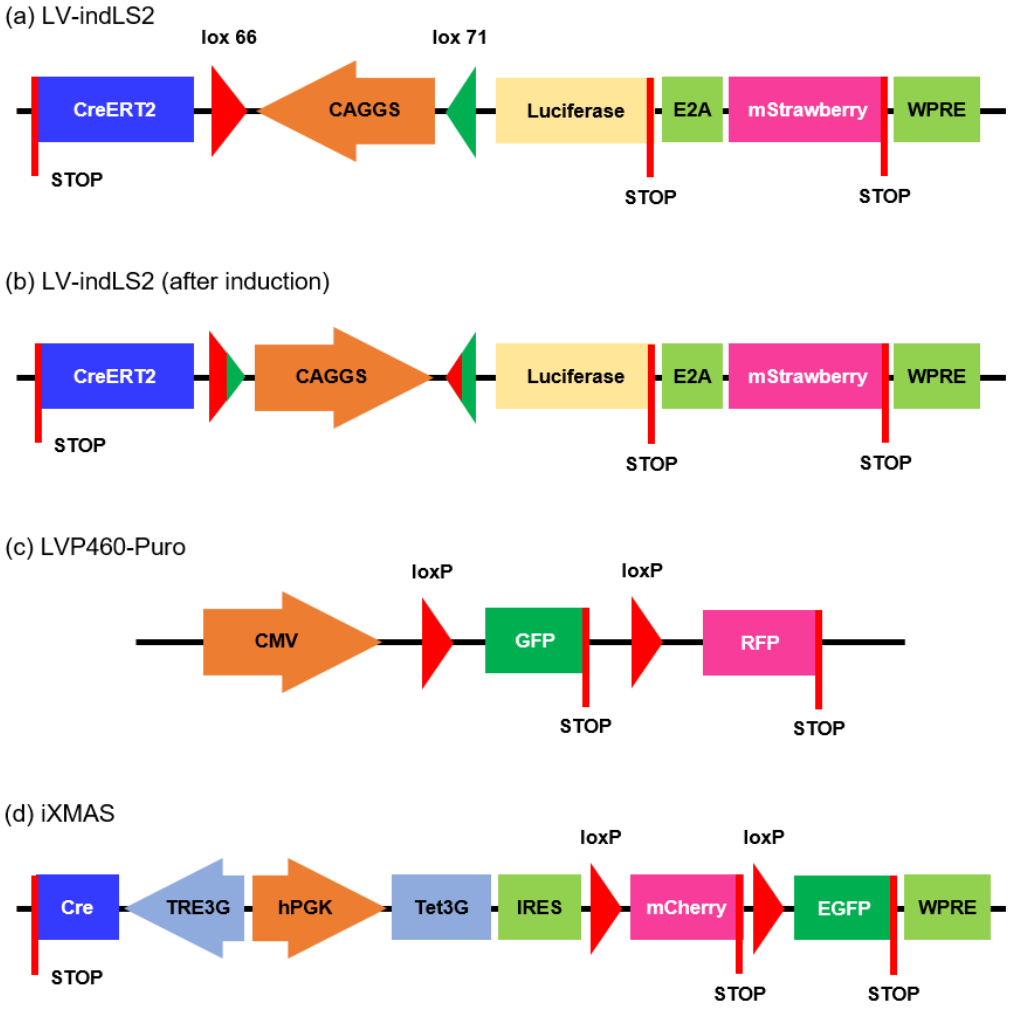


Figure 6.2: The three plasmid constructs tested for use in lineage tracing studies. The construct LV-indLS2 shown in (a) is also shown in its 4-hydroxytamoxifen induced form in (b) with the CAGGS promoter excised and reinserted in the opposite orientation. LVP460-Puro (c) is shown prior to induction with Cre, and iXMAS (d) is shown prior to induction with doxycycline. When these latter two constructs are induced, Cre excises the element found between the two identical loxP sites.

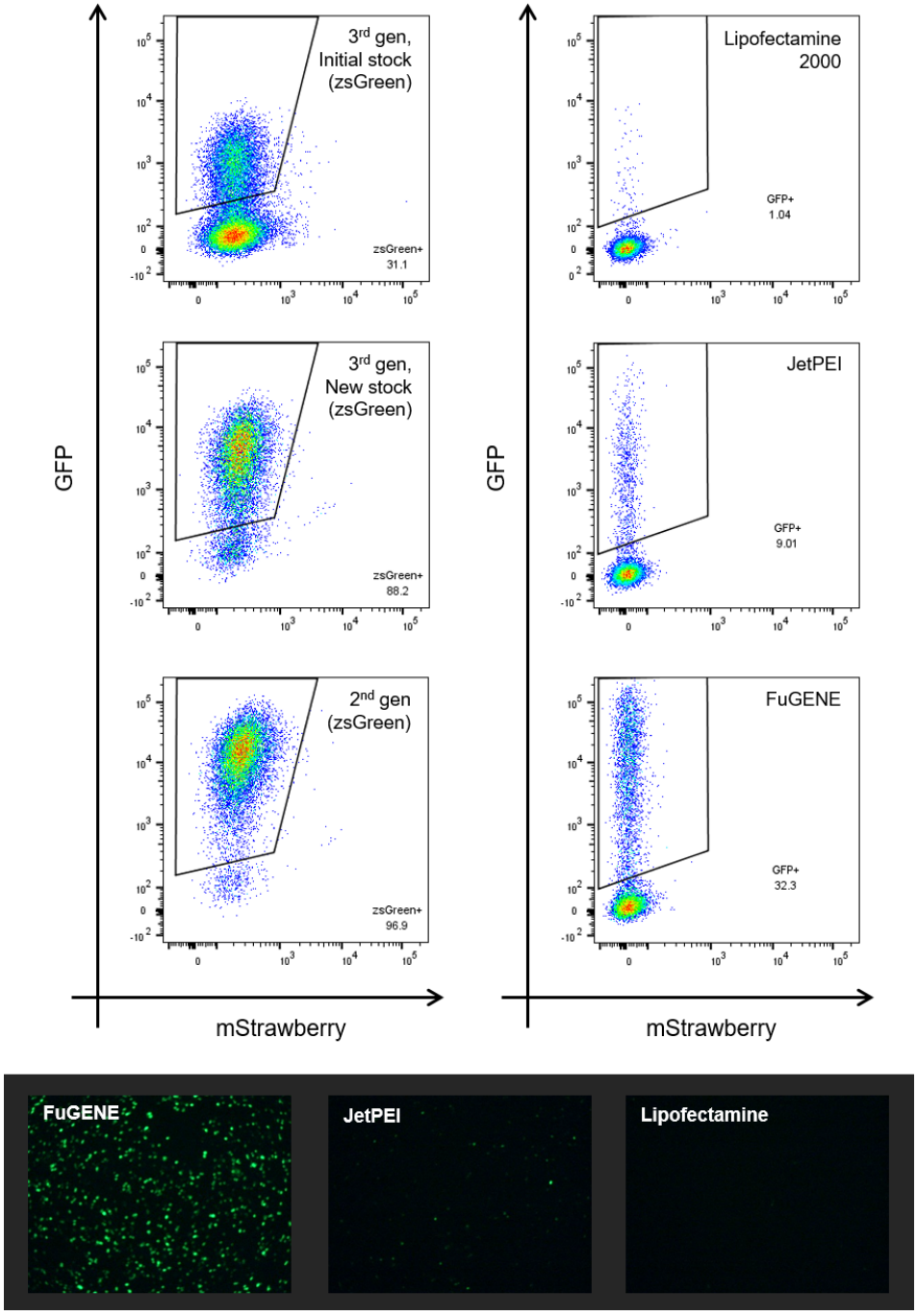


Figure 6.3: Side-by-side comparisons of the packaging and transfection types used during viral work. Left column: zsGreen control construct packaged and transduced under the same conditions using jetPEI and (top) initial sub-par 3rdgeneration packaging, (middle) replacement 3rdgeneration packaging, (bottom) 2ndgeneration packaging. Right column: Comparison of transfection reagents by parallel transient transfection of zsGreen control construct. An excess of cells was used to ensure sensitivity (since this is a very easy to transfect plasmid). Bottom panels: fluorescence microscopy of the cultures used for this flow experiment. All experiments performed in 293T cells.

cells should increase the efficiency of virus generation. A 2nd generation packaging construct (bottom left; kindly provided by Greg Hannon's laboratory) was tested against the Brenton Laboratory's third generation packaging system (top left). This resulted in a drastic increase in efficiency, that is higher than would be expected from simply switching between the systems. This suggests previous poor results might be down to an issue with the previous packaging system. When a replacement stock of the same 3rd generation packaging was tested, it showed drastically better results (although not as good as the 2nd generation), suggesting the previous packaging construct DNA may have been defective or degraded. As the 2nd generation packaging gave the best results, this method was carried forward.

A wide range of transfection methodologies were also tested during early work on this part of the project. Initially the transfection reagent lipofectamine 2000 (Thermo Fisher Scientific) was used (top right). Alternative reagents jetPEI (Polyplus transfection) and peqFECT (Peqlab) were tested as alternatives. PeqFECT appeared less efficient, but jetPEI showed increased efficiency (middle right) and less toxicity, and was used in subsequent experiments. Calcium phosphate co-precipitation was tested against jetPEI, and proved able to match its maximum efficiency. However, this method appeared highly variable in its degree of success, and was abandoned as it was hard to have confidence in results generated this way. Electroporation (using the Neon system; Thermo Fisher Scientific) was tested, and gave high transfection efficiency, but also resulted in rates of cell death >90%. This creates a number of additional concerns, primarily a severe risk of selective attrition of certain cell population (either phenotypic or at the genomic level). This risk outweighs any possible improvements electroporation might offer. Finally the transfections reagents ViaFect and FuGENE (Promega) were tested. ViaFect showed lower efficiency than jetPEI, but FuGENE consistently outperformed the previous jetPEI protocol, and was subsequently used for all following transfection experiments. The right-hand panels of Figure 6.3 show a side-by-side comparison of the three methodologies in which the constitutive zsGreen control construct was transiently transfected into 293T cells. An excess of cells was deliberately used to ensure sensitivity (hence the relatively low transfection rates). Images of the same cells, taken using a simple fluorescence microscope, are shown in the bottom panels.

Concentrations of both packaging and target DNA were optimised, and a finalised protocol developed, which is detailed in section 2.5. Figure 6.4 shows the LV-indLS2 construct tested, along with the zsGreen control and an untransduced sample, near the end of this optimisation process.

These results show effective transduction of the control construct (demonstrating that any methodological issues have been eliminated), but poor transduction of the LV-indLS2 construct, suggesting that something about this construct reduces its efficiency. The construct design has been used in previously published work, so it can be assumed to be functional. One possibility is simply that its size (14.4 kb versus 7.7 kb for the zsGreen control vector) makes viral packaging less efficient, while another is that the plasmid (or rather the stock available) itself may be defective. These hypotheses were tested in parallel by Sanger se-

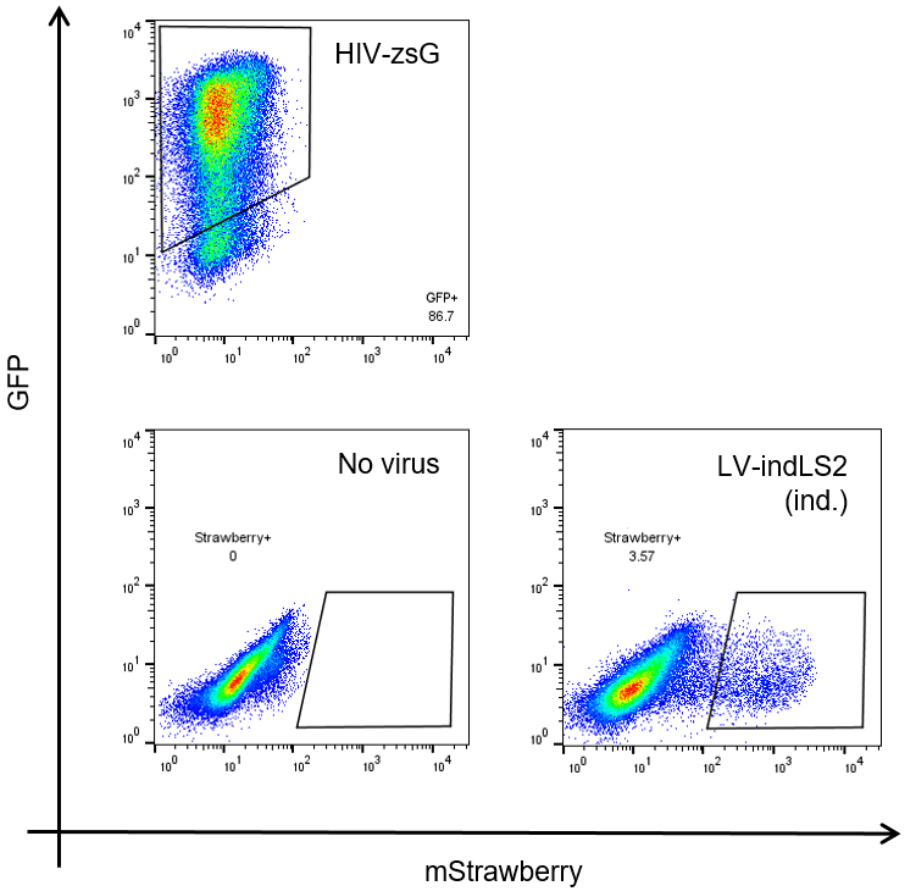


Figure 6.4: Comparison of viral generation and transduction of zsGreen and LV-indLS2, along with untransduced target cells. Transfection and transduction both performed in 293Ts. Transfection in virus generating cells was conducted with jetPEI.

quencing of key regions of the LV-indLS2 plasmid (and comparison of these to its plasmid map) and by performing transfection and transduction using an earlier generation of the plasmid (LV-LS1, a 12.2 kb non-inducible construct without the Cre element and with a constitutive CAGGS promoter driving mStrawberry and luciferase). Transduction of this construct (20%; Figure 6.5, top) was not as good as that of the zsGreen control, but significantly better than with the inducible version. Increased size impeding efficient packaging would be a consistent explanation for the reduced transduction across increasingly large constructs.

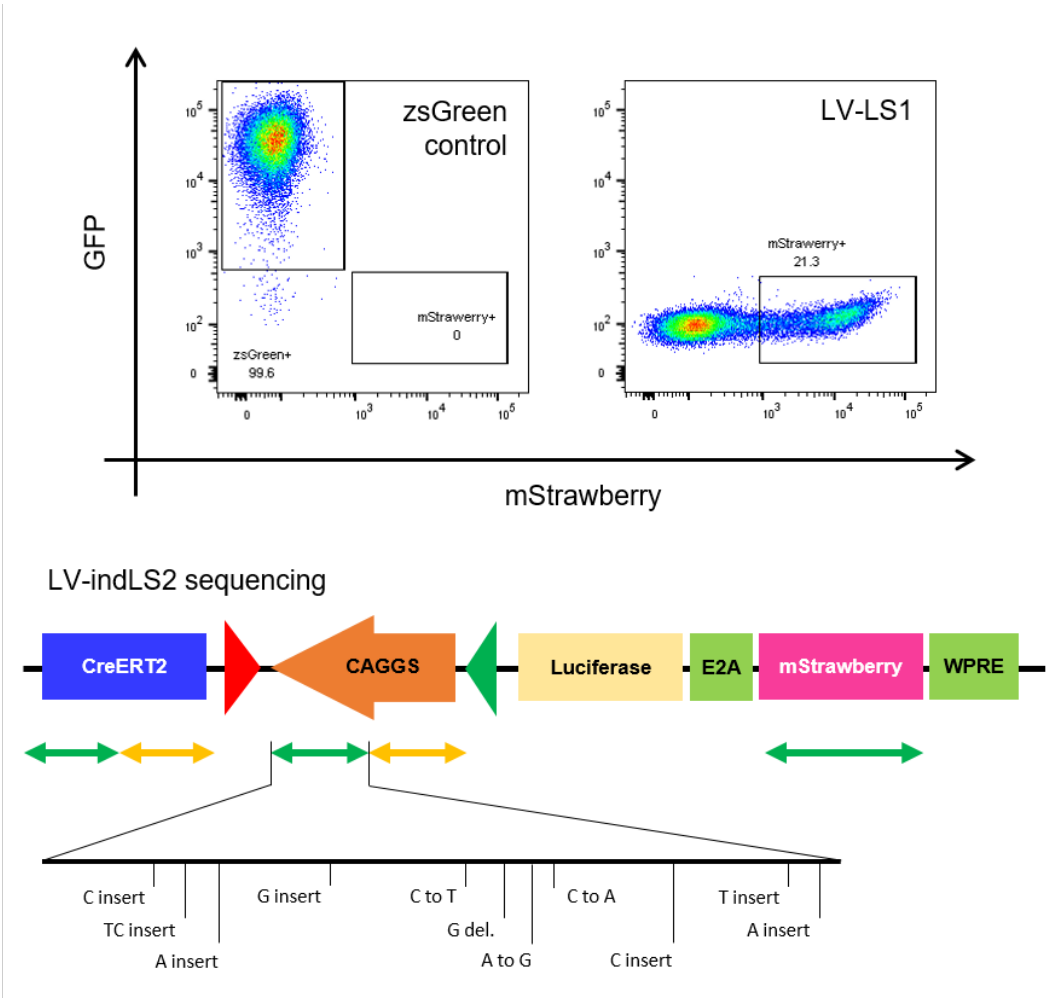
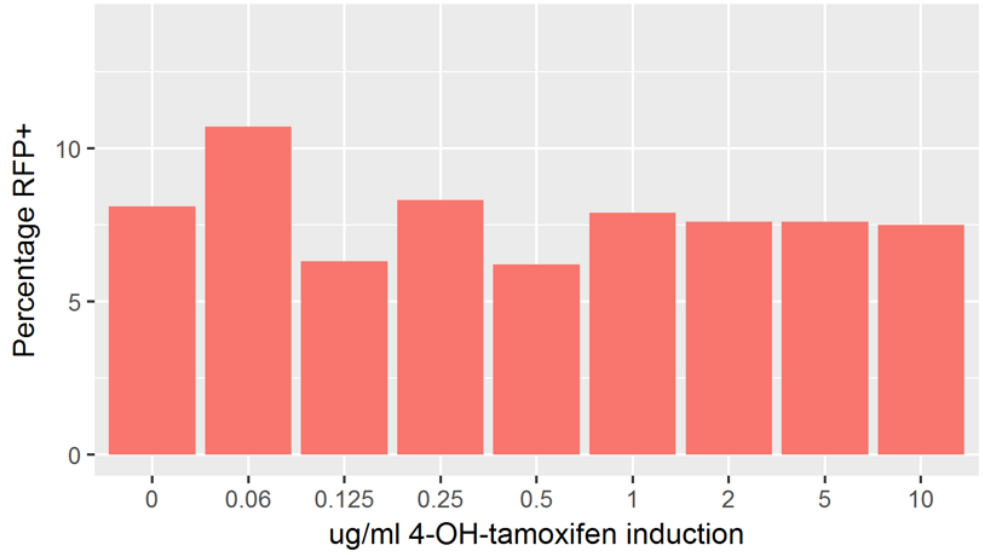


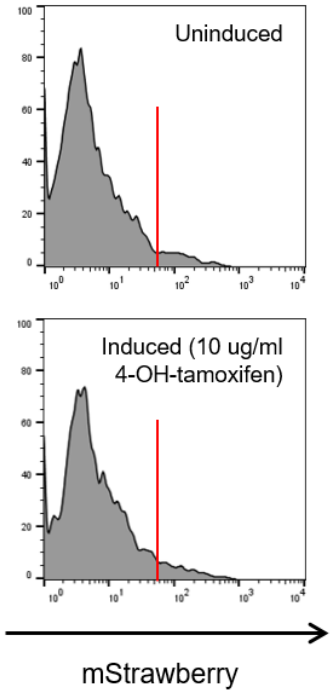
Figure 6.5: Top panel shows parallel transductions of the zsGreen control and LV-LS1 (the constitutive predecessor to LV-indLS2). The resulting transduction efficiency is much higher than that of the inducible construct shown previously (although below that of the zsGreen control). Bottom panel shows Sanger sequencing (double-headed arrows) across key regions of construct (yellow arrows show sequencing in which primers failed), and a summary of the dense region of sequence variation seen in the CAGGS promoter.

However, sequencing of key parts of the plasmid also shows a dense cluster of sequence changes in the first half of the CAGGS promoter, compared to the plasmid map (Figure 6.5, bottom). This 500 bp sequence contains 11 small sequence changes, while 3 other minor sequence changes were also noted across the remaining 2.5 kb of sequence. Such a dense cluster of variations in one area suggests a very serious risk that this sequence may not operate as expected. Since these errors are in this case found in the key promoter responsible for driving both the constitutive (Cre) and induced arms of the plasmid, this could have a very serious impact on the plasmids function. A more rigorous sequencing of the entire plasmid was not performed due to the extent of the defects found by this cursory assessment.

Titration of the 4-hydroxytamoxifen inducer (Figure 6.6) reveals yet another concern. Addition of between 0 and 10 $\mu g\,ml^{-1}$ 4-hydroxytamoxifen resulted in no change in the resulting mStrawberry expression, either in percentage of cells marked, or strength of the expression. The bottom panel of this figure shows the induced and uninduced samples against the figures validating the virus in the original paper. In Rodriguez *et al.*, the unifected cells are seen below $10^{1.5}$, the uninduced between $10^{1.5}$ and 10^3, and the induced above 10^3. In the data shown from this work, a similar pattern is seen in the uninduced sample, but there is no elevation above this baseline in the induced sample, with no meaningful cell population above 10^3 in either sample. This suggests a failure of induced expression in the induced sample. This matches the previous LV-indLS2 data seen in Figure 6.4 (where the positive population is again weakly stained, consistent with uninduced expression), but interestingly the LV-LS1 constitutive sample in Figure 6.5 shows a clear population around 10^4.

These observations all point to a failure of proper induction in the Laboratory's stocks of LV-indLS2. It also highlights the degree of uninduced 'leaky' expression of mStrawberry being produced in transfected cells since as induction does not appear to be functional, all of the visible mStrawberry is presumably leaky expression. Clonal analysis of lineage-traced tissue is dependent on being able to distinguish isolated patches of marked cells. In a tissue generating this degree of leaky expression, it will be hard not only to distinguish isolated 'islands' of truly induced clonal patches amid the background, but will also be hard to distinguish the exact boundary of the patch if it is surrounded by cells showing leaking expression. Mis-assignment of a few cells from one subpopulation showing leaky expression into a clonal patch of another subpopulation would provide false evidence for a hierarchy which may not exist. Since the nucleoside analogue tracing studies presented previously (chapter 5) suggest that if this process occurs it is quite rare, accumulation of these aberrant assignments could quickly result in the wrong conclusions being drawn from the lineage trace study. In addition, if the $EpCAM^{+}$ and $EpCAM^{DIM}$ subpopulations show differences in expression patterns, it is possible that they may impose some degree of regulation on the transduced mStrawberry gene. If this occurs, it is quite possible that uninduced expression from a high-expressing subpopulation might appear equivalent to the induced cells of the other. Equally induced cells of the low-expression subpopulation may be discarded in error if they appear to similar to the backgroud leakiness. For these reasons the degree of inherent leakiness seen from LV-indLS2 is concerning, and makes use of this





Redacted for copyright reasons

Adapted from Fig. 1d Rodriguez et al. 2014

Rodriguez, E.; Mannion, L.; D'Santos, P.; Griffiths, M.; Arends, M. J.; Brindle, K. M. & Lyons, S. K. Versatile and enhanced tumour modelling in mice via somatic cell transduction The Journal of Pathology, Wiley-Blackwell, 2014, 232, 449-457

Figure 6.6: Percentage of cells expressing mStrawberry appears to be inducer-independent (top panel). There is no meaningful difference between the flow profiles of induced and uninduced cells, and the positive cells are only weakly red, not exceeding 10^3 (bottom left). Figures from the original paper (bottom right) show their uninduced cells having a similar pattern of weak expression, suggesting that no induction is being seen in this case.

construct for this type of study a risky endeavour.

Alternative construct options were considered. The construct LVP460-Puro (see Figure 6.2 (c)) is sold commercially by Amsbio as a pre-packaged virus, and appeared to be ideal for use in this project. This ~10 kb construct constitutively expresses GFP, until induced by Cre. Cre excises the GFP gene, allowing the constitutive promoter to instead drive RFP. Given the apparent inefficiency of large constructs, the relative simplicity of this lentivirus is a major advantage to the design. The absence of the Cre gene from the virus itself means it must be inserted by other means. This adds a degree of additional complexity, but also drastically reduces the chance of unintended recombination since there should be no risk of low level expression of the Cre protein which can cause this. The inclusion of a constitutively expressed fluorophore provides several advantages, in particular allowing for the efficiency of transduction to be more easily quantified. Additionally, using two separate fluorescent genes allows promoter-independent expression of the inducible fluorophore to be distinguished from true leakiness (uninduced recombination leading to constitutive expression of the inducible gene), as the former would result in weak and transient expression of the inducible in addition to the constitutive, while the latter would show strong expression of the inducible, and loss of expression in the original consitutive gene (as recombination removes this cassette entirely).

Pre-packaged virus of this design was tested on an ovarian cell line (OVCAR-3) and a dissociated xenograft sample (JB215) at an MOI of 2 (200k viral units added to 100k cells in a 24-well plate). In addition to testing uninduced construct, induction was tested by transduction with a simple constitutive virus expressing Cre and GFP. The results were disappointing. No fluorescence was seen for any of the uninduced samples (transduced with LVP460-Puro alone, which in theory should be GFP^+). As the construct is designed to switch from green to red upon induction, a successful induction (samples transduced with LVP460-Puro and Cre-GFP) should be indicated by the appearance of RFP signal amid the GFP. Instead, there was a shift from no signal to GFP positivity, consistent with successful expression of the Cre-GFP inducer virus, but no expression from the LVP460-Puro. The induced samples were indistinguishable from controls infected with the Cre-GFP construct alone.

Further commercial options were assessed, but it proved challenging to find a suitable inducible construct which met both key criteria: small size and strict induction (minimal leakiness). Eventually it became clear that the best option might be to generate a new construct purpose built for this task. In collaboration with the CRUK Cambridge Institute Genome Editing Core, a design was developed and synthesised using the VectorBuilder tool available from Cyagen. The construct (along with pre-packaged virus) were then synthesised by Cyagen.

The resulting construct, iXMAS (Figure 6.2 (d)), constitutively expresses tet3G and mCherry, while uninduced. Transcription of these two proteins is from the same constitutive hPGK promoter, with an IRES site allowing for initiation of translation mid strand, allowing both proteins to be generated from the same promoter. Upon doxycycline induction, doxycycline

binds to Tet3G protein allowing it to in turn bind at the Tet-response element (TRE3G). Tet3G binding to TRE3G activates this promoter and initiates expression of Cre. Cre protein then cleaves the insert at the loxP sites, allowing recombination and excision of the mCherry gene. This effectively replaces the mCherry gene with an EGFP gene, resulting in a red-to-green colour switch when induced. Given the apparent inefficiency of large constructs, the relatively small size of this lentivirus (11.1 kb versus 14.4 kb for LVindLS2, and 7.7 kb for the zsGreen control) is a major advantage to the design.

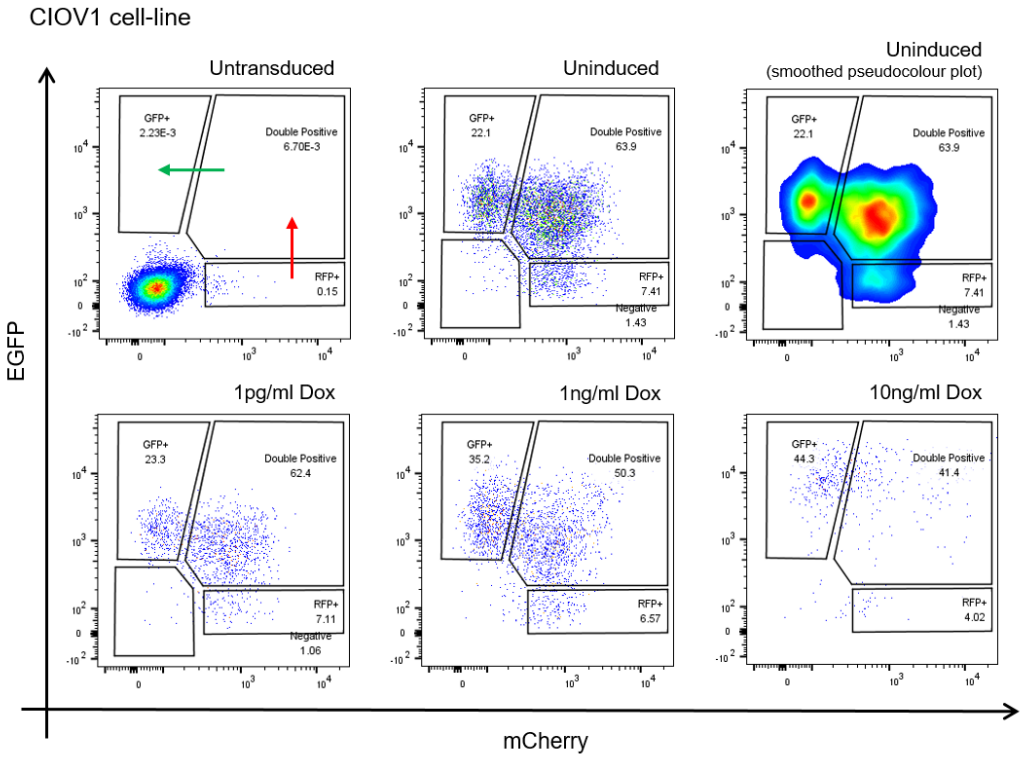


Figure 6.7: Flow cytometry profiles for cells infected with iXMAS. The arrows on the first panel (control cells untreated with virus) indicate the theoretical shift from RFP^+ to GFP^+ expected upon induction. Cells from the CIOV1 ovarian cell-line were transduced at an MOI of 5, and where indicated dosed with doxycycline (Dox). The third figure in the top panel duplicates the data in the previous figure using smoothed plotting to show density.

The pre-packaged iXMAS virus was applied at an MOI of 5 (10^7 titred viral units to 200k target cells). The HGSOC cell-lines CIOV1 and CIOV3 (developed in the Brenton Laboratory by Maria Vias; manuscript in preparation) were used as target cells for the transduction. Pilot work with the constitutive control viruses shows these cells are harder to transduce than 293Ts, but easier than cells fresh from patients. Interestingly passaged xenografts appear to have a similar transduction efficiency to that of these cell-lines, rather than that of patient material. This might indicate subtle changes in these cells as a result of their

in vivo passage, but may also just indicate that transduction of patient cell is hard due to factors beyond the cells themselves, such as cells fresh from patients being surrounded by non-tumour cells (e.g. red blood cells) which may soak up virus. In addition, passaged xenograft cells may be more resilient to *in vitro* culturing as a result of their exposure to the xenograft environment, whereas cells fresh from the patient may be stressed by the transfer *in vitro* resulting in many infected cells undergoing apoptosis (or at least not growing) as a result of the additional stress of viral transduction.

As shown in Figure 6.7, transduction with the pre-packaged iXMAS virus without induction resulted in extensive fluorescence. It would be expected that this experiment would show only mCherry signal, since without induction there should be no expression from the EGFP gene. While ∼80% of cells do indeed show mCherry signal (an encouraging degree of uptake), ∼90% also show EGFP signal suggesting that there is either very high promoter-independent expression from this gene, or that there is unintended induction of the construct occurring in the absence of inducer addition. Addition of varying concentrations of Doxycycline (from $1\,\mathrm{pg\,ml^{-1}}$ to $10\,\mathrm{ng\,ml^{-1}}$ final concentration) reveals that with increasing concentration, there appears to be a shift from ∼90% $\mathrm{EGFP^+}$ to ∼95% $\mathrm{EGFP^+}$ at the highest concentration. There also appears to be significant toxicity at the higher doxycycline concentrations, with cell numbers drastically reduced, and it is quite possible the shift in populations is entirely down to relative resilience to the toxicity.

The presence of $\mathrm{mCherry^+}$ cells in the induced samples is not necessarily a major concern, as even correctly induced cells may retain some of their previously expressed mCherry protein. However, the presence of strong EGFP signal in the uninduced sample was unanticipated and raises serious questions about the viability of this construct for inducible lineage tracing.

That the uninduced sample appears to contain distinct $\mathrm{mCherry^+/EGFP^-}$ and $\mathrm{mCherry^+/EGFP^+}$ subpopulations (see smoothed pseudocolour plot) would suggest that this is unlikely to be weak promoter-independent expression since this would be expected to create a random spread rather than a distinct bimodal distribution (which is more indicative of a binary between recombined and non-recombined). However, there is also a noticeable shift in the magnitude of EGFP expression (from $\sim 10^3$ in the uninduced to $\sim 10^4$ with $10\,\mathrm{ng\,ml^{-1}}$ doxycycline) suggesting that strong induction is causing a meaningful increase in signal. This would imply that the uninduced expression may in fact be weak promoter-independent, since if aberrant recombination were happening no effect should be seen on subsequent induction, whereas a contrast would be expected between weak promoter-independent expression and induction-induced promoter-dependent expression.

As shown in Figure 6.8, similar patterns are seen when xenograft material is transduced. The rate of fluorescence is much lower here primarily because a much less concentrated titre was used (as this was simply a proof-of-concept to determine if similar aberrant induction was seen). However, it is immediately clear that the relatively weak fluorescence from this construct is problematic in this context, as xenograft cells are much less uniform than cell-lines, and the fluorescence profile of untransduced control cells is already quite broad. As

such, it is hard to reliably distinguish genuine fluorescence from the autofluorescent 'tail' of untransduced cells.

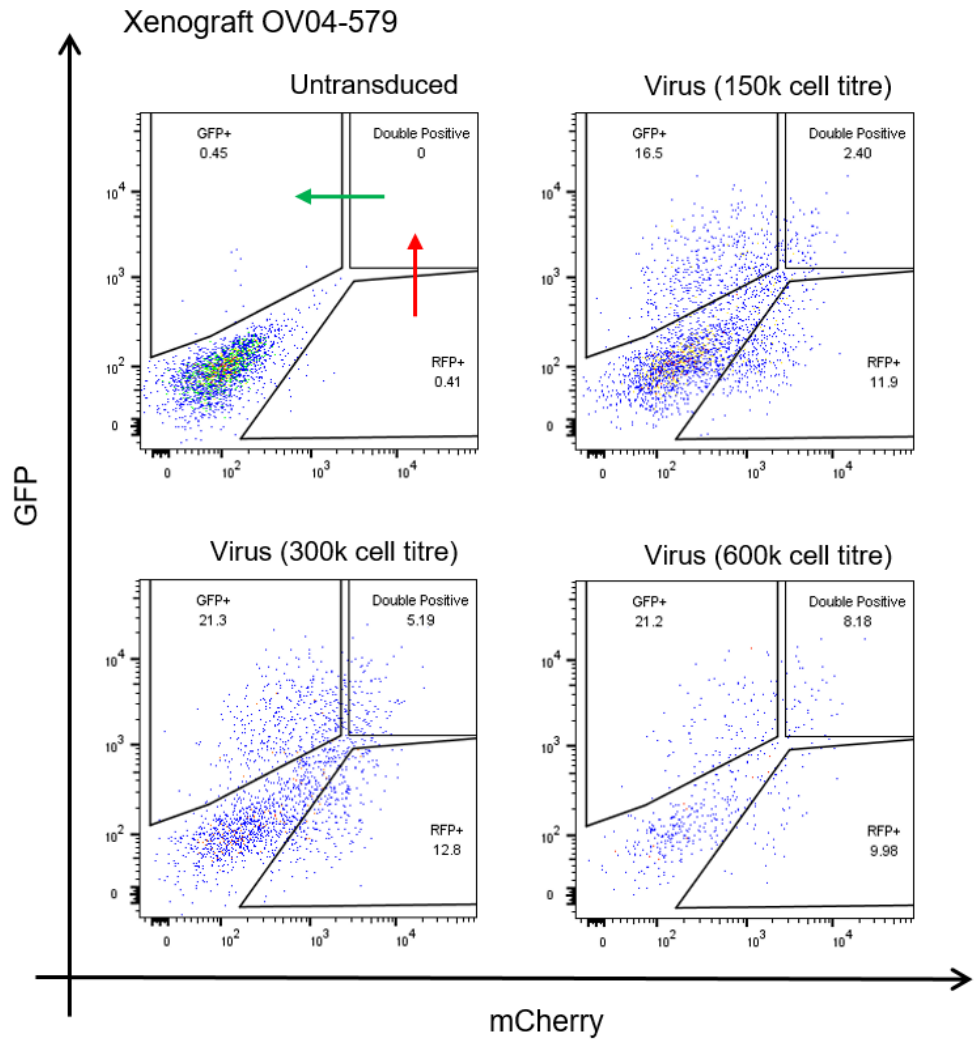


Figure 6.8: Transduction *in vitro* of cells from a patient sample (versus untransfected) shows that it is even harder to rigorously detect fluorescence positive cells in this context due to the noisier baseline. The toxicity at highest titre is suspected to be an isolated technical issue.

In order to determine the cause of the cryptic results from these experiments, the entire insert of the iXMAS construct was sequenced (from Cre through to WPRE). The results showed an entirely pristine sequence across this entire region. Western blots were then performed to assess the expression of the fluorescent proteins, and determine if the multicistronic cassette was operating correctly. As seen in the left-hand panel of Figure 6.9, the uninduced

samples show good expression of mCherry, while this protein is largely eliminated in the induced samples. This shows that induction is correctly excising the mCherry cassette upon induction. However, this panel also shows that there is a marked increase in Tet3G expression upon induction (bottom-left panel). This is not expected since the Tet3G gene is directly in front of the constitutive promoter and before the IRES and loxP sites. It is unclear how this effect is arising, but a plausible explanation would be that the hPGK promoter (driving Tet3G and the fluorophores) is being influenced by changes at the neighbouring TRE3G inducible promoter. When doxycycline addition causes Tet3G protein to bind to the TRE3G inducible promoter and activate Cre, it is possible that it causes sufficient conformational changes in the DNA that it is able to make transcription from the hPGK easier, leading to enhanced expression of the theoretically constitutive proteins.

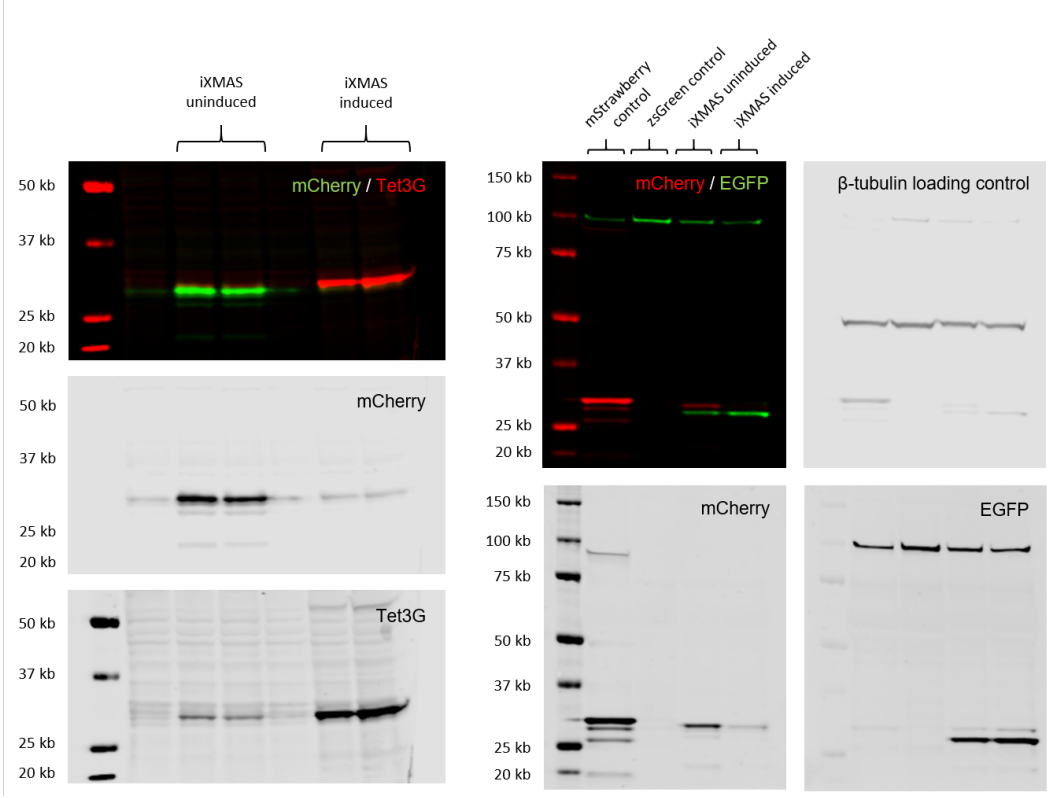


Figure 6.9: Western blots to assess Tet3G, mCherry, and EGFP expression. Left and Right panels are distinct experiments. BioRad All-Blue ladder is used in both cases; molecular masses indicated. Lanes 1 and 4 of left-hand figure are blank and loaded with water.

This might also explain the apparently contradictory results described previously. If induction and Tet3G binding results in an increase in transcription from the constitutive promoter (as Tet3G expression suggests) then it is possible that the same could be happening at the IRES site despite the relative distance. If this is the case, then induction should not only cause the mCherry cassette to be replaced by EGFP and upregulate Tet3G, but also cause

an upregulation of EGFP expression (even in cases where the loxP excision has already occurred). Thus large concentrations of doxycycline would lead to amplification of $EGFP^+$ expression even though most cells had already aberrantly recombined. Further investigation would require separation of the effects of Tet3G binding and induction. Testing the effects of induction using Cre expressed by a separate construct would be a logical step for further investigation of this (however the only constitutive Cre easily available in this case co-expresses GFP and is therefore not helpful).

The right-hand panel shows comparison of the expression of mCherry and EGFP. Both these proteins are around 25 kDa in mass, and it is clear that while mCherry signal is restricted to the uninduced construct in lane 3 (plus the constitutive mStrawberry control in lane 1), EGFP signal is seen in both induced and uninduced samples (lanes 3 and 4). This indicates that there is significant expression of GFP prior to induction. The zsGreen control (lane 2) does not show any signal as zsGreen and EGFP do not share homology (unlike mStrawberry and mCherry which are variants of the same RFP sequence). This sample was included as a negative control to exclude off-target detection, for example the 100 kDa band seen in all samples which appears to be the EGFP antibody binding to a ubiquitous target. A β-tubulin loading control is also shown to confirm similar concentrations of protein were loaded.

Finally, Figure 6.10 shows images of the cells used for these western blots taken under a fluorescent microscope (Nikon TS-100 with pE-300^{white} illuminator). It is clear from this that although fluorescent protein from iXMAS can easily be detected by western blot, the concentrations are such that it is very hard to reliably detect via fluorescence microscopy. It also clear that it is not possible to detect a difference in EGFP expression between induced and uninduced by this method at the level of field of view, let alone at the level of single cells. While there are ways around the sensitivity problem (for example IF or IHC staining could be performed for GFP and RFP on fixed slides), the inability to meaningfully distinguish induced cells from uninduced ones at a single-cell level means that this construct is not viable for inducible lineage tracing.

6.3.2 Symmetry assays

PDX tumours from 7 of the established and validated PDXs were pulse chased with 100 mg kg^{-1} BrdU and mice killed and tumours extracted and formalin fixed 6 h later. The initial methodological plan for this experiment was to cut absolute serial sections, and IHC stain for BrdU and EpCAM using the established methods available through the CRUK Cambridge Institute Histopathology Core pipeline. To initially confirm the presence of the markers of interest, single sections were cut, and separately stained for BrdU and EpCAM, along with p53 to confirm tumour status. This pilot work showed good staining for BrdU (such that pairs could be efficiently distinguished) and p53 confirmed these cells were in fact tumour. However, while EpCAM staining was adequate to distinguish epithelial cells from stroma,

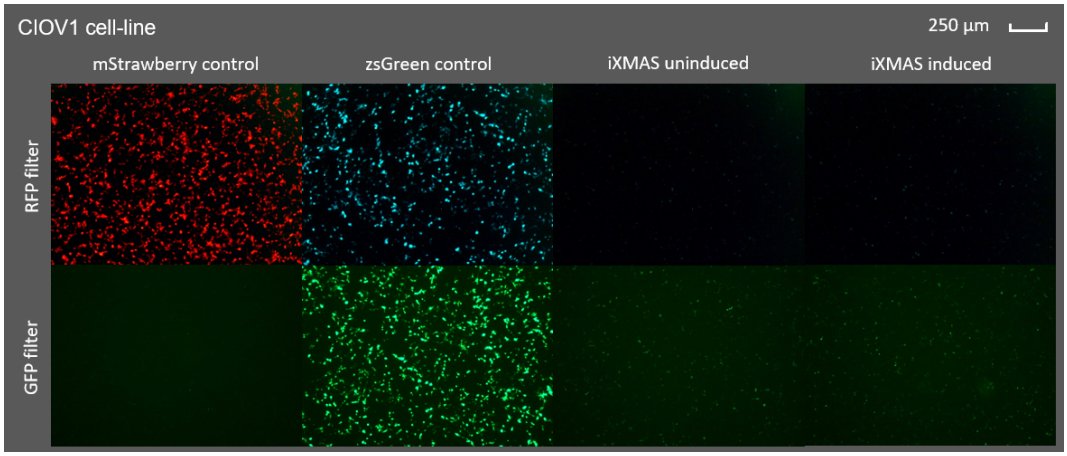


Figure 6.10: Images of cells transduced with control constructs and induced and uninduced iXMAS, taken using a simple fluorescence microscope. iXMAS expression of fluorescent proteins is significantly below that of the control constructs. Induced and uninduced cells cannot be practically distinguished.

it quickly became clear that this methodology was not able to quantitively distinguish between $EpCAM^{+}$ and $EpCAM^{DIM}$ cells, as both would appear positive. This suggests that the standard IHC stain available from the CRUK Cambridge Institute Histopathology Core has been optimised with a view to maximum sensitivity, as it is able to stain even weakly stained cells well. However, in ovarian epithelium where EpCAM expression is very strong, it ceases to be effectively quantitative because all cells are picking up high degrees of stain. In addition, p53 staining showed a number of non-tumour cells within what was assumed to be solely tumour. This raised further issues as the analysis software used by the Core – HALO (Indica Labs) – is able to map two serial sections (enough for BrdU and EpCAM) but not three (as would be needed in order stain p53 as well, in order to confirm HGSOC status at the cellular level).

To circumvent these issues, an immunofluorescence (IF) approach was adopted instead. IF allows for a more quantitative assessment of the presence of EpCAM, which should be sufficient to distinguish $EpCAM^{+}$ and $EpCAM^{DIM}$ subpopulations. Additionally, by using an IF approach, a single slide can be triple-stained for EpCAM, BrdU and p53, eliminating the complexity of mapping multiple serial sections, and eliminating the issues with mapping a third serial layer. However, since this is not an established technique, nothing beyond the preparation of blank slides was available through the Core.

Methods for staining BrdU by IF had already been optimised as part of pilot work for the nucleoside tracing experiments described previously, while p53 IF was optimised for validation shown in Figure 4.9. The former used an acid retrieval method (2 M HCl for 30 min at room temperature), but BrdU proved to be detected just as well when treated with the citrate method used for the latter (described in subsection 2.4.1). Detection of EpCAM proved significantly harder. Several EpCAM antibodies were tested (see Table 2.3) using citrate antigen retrieval (as above) for 10, 20, 30 and 40 minutes, HCl retrieval, and no retrieval.

None of these methods gave any staining over and above that from secondary-antibody-only controls. Various additional common methods were subsequently tried including EDTA (Ethylenediaminetetraacetic Acid) retrieval (boiling for 10 min in 1 mM EDTA and 10 mM Tris at pH 7.4), alkaline EDTA retrieval (boiling for 10 min in 0.1 mM EDTA and 10 mM Tris at pH 9), and proteinase K (PK) retrieval (20 µg ml^{-1} proteinase K (Roche Diagnostics) in TE Buffer for 20 min 37 °C).

The PK retrieval in combination with the EpCAM antibody VU-1D9 (AbCAM) gave good strong signal localised to the cell surface, and present only in cells of an epithelial nature. Furthermore there was variation in signal strength between different cells, suggesting that the method has the sensitivity to separate EpCAM^{+} from EpCAM$^{\mathrm{DIM}}$ cells.

In order to stain for both EpCAM and p53 on the same slide, combined antigen retrieval methods were tested. Both acid retrieval and citrate retrieval (as previously described) were tested in combination with the successful PK digestion (in both possible orderings). PK followed by either acid or citrate retrieval resulted in complete cellular disintegration. When either acid or citrate was followed by PK retrieval, tissues remained fully intact, but no EpCAM staining was observed (in contrast to the PK-only control performed in parallel which showed good staining). Thus is does not appear to be possible to dual stain the same slide for both EpCAM and p53 using any of the common antigen retrieval methods.

As such, it appears triple-staining is impractical without starting afresh with untested antibodies. One possible approach – and the probable next step in development of this method – would be to perform separate dual- and single-stainings on absolute serial sections and use the HALO system to merge the images. The specific antigen retrieval of the EpCAM antibody make it a logical choice for the single-stain. One of the biggest challenges of triple-staining is that both the antibodies seen to work well in these pilot experiments (DO-7 for p53, and VU-1D9 for EpCAM) are mouse-derived, and working around this (either with new antibodies, or secondary detection specific to isotype) adds complications. Switching to combined dual / single stains allows both of these optimised antibodies to be used on separate slides with the same secondary antibody. Previous pilot work in this project (not discussed) involved optimisation of several antibodies for BrdU. These are all mouse or rat derived (and the latter has shown cross-reactivity with mouse secondaries, even cross-absorbed types). However, the CRUK Cambridge Institute Histology Core have optimised a sheep-derived anti-BrdU antibody for routine IHC work, using an HCl based antigen-retrieval similar to the one used successfully with DO-7 for p53. Therefore, it is possible these two antibodies could work well together with minimal method-development.

Unfortunately, further development of this methodology is not possible within the timeframe of this project.

6.3.3 Single-cell mRNA Sequencing

PDXs derived from 8 different patient samples were sorted. Before dissociation, a small piece of each xenograft was cut and frozen for TAm-Seq to make absolutely sure these key xenografts retain the same mutational 'fingerprint' as the original tumours, and early xenograft passages. All results were as expected with no changes observed; the key data are summarised in Table 6.1, and confirm the identity of the specific xenograft tumours being used for subsequent scRNA-Seq.

For each PDX sample, the $EpCAM^+$, and $EpCAM^-$ subpopulations were isolated using flow sorting (FACS Aria). An unsorted sample, not subject to flow sorting, was also collected. In 4 of the 8 samples a $CD90^-$ subpopulation was also observed, and this population was sorted too.

Xenograft validation by TAm-Seq			
PDX	PDX *TP53* mut.	Patient	Patient tumour *TP53* mut.
OV04-496 xenograft	c.1024C>T	OV04-496	c.1024C>T
OV04-581 xenograft	c.743G>A	OV04-581	c.743G>A
OV04-516 xenograft	c.742C>T	OV04-516	c.742C>T
OV04-072 xenograft	c.574C>T	OV04-072	c.574C>T
OV04-449 xenograft	c.403T>G	OV04-449	c.403T>G
OV04-366 xenograft	c.783delT	OV04-366	c.783delT
OV04-348 xenograft	c.637C>T	OV04-348	c.637C>T
OV04-075 xenograft	c.991C>T	OV04-075	c.991C>T

Table 6.1: Summary of the results of TAm-Seq performed on the xenograft samples used for scRNA-Seq. The fingerprint of a single key mutation in p53 is conserved between these samples and the patient samples from which they are derived.

After library preparation as described, the libraries were pooled, normalising based on total cDNA concentration, and a single lane of sequencing performed. The resulting data were analysed, in collaboration with the CRUK Cambridge Institute Bioinformatics Core, to assess the quality and quantity of each pool, with a view to ascertaining a more accurate picture of the cDNA available (for the purposes of an optimised pooling / normalisation strategy). The Cell Ranger analysis pipeline (10x Genomics) was used to demultiplex (separate each sequence read based on the sample index), assign each read to a partition (based on the partition-specific 10x barcode), to a unique cDNA strand within the library (based on the randomer barcode), and align the read to the genome. In this case since both human and mouse mRNA products are being sequenced, alignment must be attempted against both genomes. From this the Bioinformatics Core can output a 4 column *.csv* file listing partition barcode, UMI-count assigned to the human and mouse genomes respectively, and Cell Ranger's assignment of the cell based on the data (either 'Human', 'Mouse', or 'Multiplet'). Partition barcode is simply a list of the barcodes assigned to each 10x labelling bead (at

least those in droplets containing cells); this is in theory a list of cells. UMI (Unique Molecular Identifier) refers to unique cDNA elements within the cDNA library, and so the UMI is based on the randomer barcodes on each primer on the labelling bead. Effectively UMI is counting the number of mRNA strands in the original reverse-transcription (plus noise), as opposed to 'read number' which counts the number of times a particular UMI / all UMIs within a sample pool have been sequenced. Cell Ranger appears to perform some form of quality-control (QC) on the data prior to generating this file, removing empty partitions (close to 90%) and those likely to be anomalous, such as those with near-zero UMI numbers (e.g. empty partitions with a single free-floating mRNA), or near-zero read numbers (e.g. arising from PCR errors in the partition or index barcode, creating reads which get assigned to an empty partition), hence the list lacks any option for assigning cells as 'empty'.

The opacity of this QC process was the first of several issues which needed to be addressed before any serious analysis could begin. Inspection of several samples suggested that in some cases, Cell Ranger's calling is unreliable. This appears to be due to Cell Ranger not being designed with xenograft experiments in mind. For most samples the distinction between mouse and human is relatively unambiguous (for example the sample in the top-left panel of Figure 6.11), and Cell Ranger's calling is entirely reliable. The problem appears to be samples with very few mouse cells, where Cell Ranger cryptically becomes hypersensitive. For example, in one sample (OV04-366 CD90$^-$), Cell Ranger rejects ∼600 partitions as multiplets, when it is clear they are actually human cells (see Figure 6.11 top-right panel – all the cells in blue are called as multiplets, despite many on the right of the plot clearly being human). In all of these ∼600 cells, there are >1000x as many human UMIs (10k-1k) as mouse UMIs (10-1).

The Bioinformatics Core suggest this may be a result of the algorithm benchmarking against the average number of mouse UMIs in the sample. As a crude conceptual example, for a sample where half of (non-empty) partitions have 0 mouse UMIs, and the other half have 1 mouse UMI, the algorithm calculates the average mouse UMI as 0.5, and then considers everything higher (i.e. the 50% of partitions with 1 UMI) as 'mouse'. It then performs the equivalent calculation for human and then discards cells with 1000 human UMIs and 1 mouse UMI, as multiplets. The actual calculation process is more complex than this, but this hypothetical example demonstrates the apparent conceptual source of the problem. As such, this issue seems particularly pronounced in some CD90$^-$ samples as there is near-zero mouse contamination.

From the data output discussed above, it is not possible to directly examine if there is a similar calling problem in samples with very low UMI counts (i.e. that blank partitions may be called as cells, or cells may be called as blank partitions), as empty partitions are eliminated from the output. However, manual investigation of the data outputs implies that this may be the case; for example in the data for the sample discussed above, there are 36 'cells' containing 1 mouse UMI and 0 human UMIs. In most samples no partitions are shown with values this low. Since it would be reasonable to expect at least some partitions

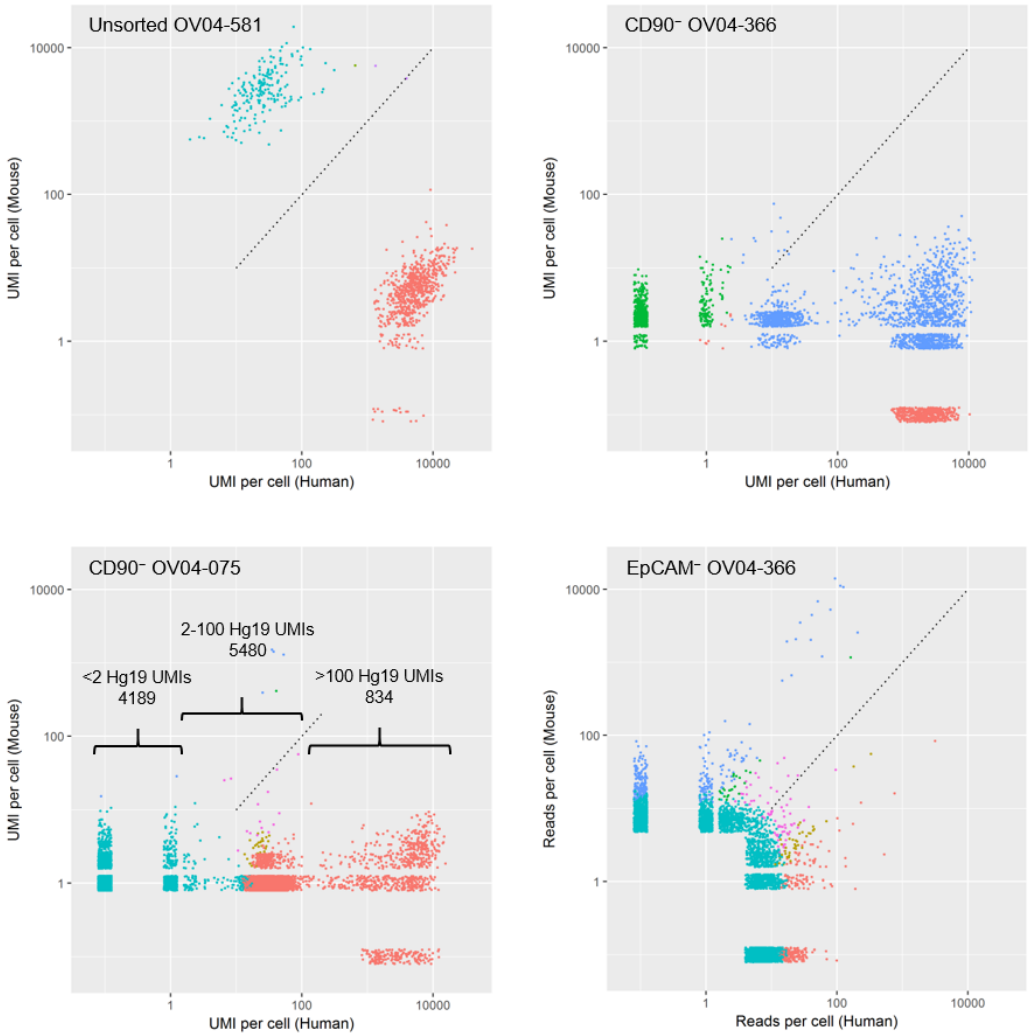


Figure 6.11: Four examples of plots showing the number of UMIs per cell (each data point is a 'cell') mapping to the human and mouse genomes. The top-left panel shows an ideal sample in which human and mouse cells are clearly separated, and calling by Cell Ranger or by inspection or my own R script produces identical results. Top right shows a sample in which Cell Ranger calls cells multiplets (data shown in blue) despite the right-hand blue cluster clearly being human cells (cells in green are called by Cell Ranger as mouse, and red as human). Bottom left shows a similar profile divided by inspection into (L-R) empty partitions, unknown group (possibly human cellular debris), human cells. Bottom-right shows a sample with too many cells which has resulted in very low saturation and thus very low UMI counts – it is not possible to determine likely split in populations from these data. The bottom two plots are coloured using my R script (with probable human cells in red, those with insufficient data in turquoise), however in the edge-cases shown this approach is questionable, and this colouring should be treated with scepticism, particularly around cells with less than 100 Hg19 UMI.

to contain single mRNA molecules (if only by PCR error on the partition barcode) it can be presumed that these partitions are normally being discarded by QC, and that their presence in this sample is another artefact of the baselining effect discussed above.

These problems can be mostly overcome by discarding Cell Ranger's calls and working directly from the UMI data from its pipeline. In order to do this, Cell Ranger was used to generate the raw UMI counts, and calling estimated using my own crude R script (based solely on ratios between human and mouse UMIs). Since Cell Ranger's QC processes do not appear to routinely eliminate real cells, but often allow dubious non-cell-partitions, or miscall them, this method should have minimal chance of resulting in the loss of true human cells. For the majority of samples, this generated a very similar output to Cell Ranger, but allowed manual inspection of the data, allowing for a greater degree of confidence in the output, and a more reproducible analysis. For a minority of samples where there was much ambiguity, this hybrid analysis method allowed for manual inspection of atypical samples, allowing sensible conclusions to be drawn.

		Cell number sequenced by scRNA-Seq			
Batch	OV04 No.	$CD90^-$	$EpCAM^{DIM}$	$EpCAM^+$	Unsorted
1	348	410	1800	535	340
1	516	N/A	2660	360	810
2	072	N/A	340	650	900
2	496	N/A	495	405	1910
3	075	10500	345	1285	1750
3	366	2870	4340	810	1410
4	449	3375	1790	2730	2390
4	581	N/A	2755	2470	2720

Table 6.2: Estimates of cell number from Cell Ranger analysis.

The second issue, shown in Table 6.2, is that there appears to be a great deal of variation in the number of cells being sequenced (this dataset comes from Cell Ranger's automated analysis). The intent was that 4500 cells were titrated into the Chromium chip such that 2500 should be correctly encapsulated for RT-PCR. In many cases far fewer cells appear to be being output in the subsequent library preparation.

Each of the 4 batches of samples were prepared on separate days, with cell number being estimated by library cDNA concentration as determined by Bioanalyser. Sequencing (from which the data in Table 6.2 was calculated) was not performed until all libraries were prepared.

The cell numbers for the unsorted samples gradually improve over successive batches. Since these samples were prepared from counted single-cell solution, this is likely to result from improved familiarity with the protocol and more accurate counting and titration. The cell numbers arising from the sorted samples are more perplexing since they were prepared by

FACS using rigorous gating to remove debris, and should theoretically contain an accurate number of cells. There is also a disparity (particularly in the first batch) in the number of cells in the $EpCAM^{DIM}$ and $EpCAM^{+}$ populations. These apparent high cell numbers for the $EpCAM^{DIM}$ subpopulation are particularly curious given that this subpopulation consistently had the lowest cDNA concentration by Bioanalyser assay, but may be indicative of overall lower expression by these cells.

Based on the Bioanalyser output from the first two batches (and the, now questionable, assumption that cDNA concentration correlated with cell number), the number of cells used in subsequent batches was increased 2-fold for the $EpCAM^{+}$ fraction, and 3-fold for the $EpCAM^{DIM}$ and $CD90^{-}$ fractions. In general, this alteration appears to have been beneficial, resulting in cell numbers much closer to the desired values, although there is still some degree of variation in the resulting cell numbers.

There are a couple of samples with very high cell number, namely the $CD90^{-}$ subpopulation of OV04-075 and the $EpCAM^{DIM}$ population of OV04-366 which both have drastically higher cell numbers than expected. Upon manual inspection, the former appears to be an artefact. As shown in the bottom-left panel of Figure 6.12, the 'cells' in this trace broadly divide into three groups. There are (a) $\sim$4000 partitions with fewer than 2 human reads and fewer than 10 mouse reads, which are almost certainly just the result of noise, and contain no cells, (b) $\sim$800 cells which have >100 UMIs and are almost all clearly human cells, despite being called by Cell Ranger as multiplets, and (c) $\sim$5000 partitions with between 2 and 100 UMIs. It is unclear what this population represents, since it is not intuitively obvious what would give rise to this trimodal distribution. If these partitions represent cells, then why are so many fewer UMIs being sequenced for these partitions as opposed to those getting thousands of UMIs? If these are simply artefacts, why do they have such a density of human UMIs and yet so few mouse? The most plausible hypothesis is perhaps that they represent some form of human cell debris, possibly vesicular in nature that still contains a small number of human mRNAs. If they do represent real cells (and I would argue that they probably do not), this would raise more fundamental questions about the assay, in that it would suggest that some cells are sequenced much more deeply than others. This would imply a significant bias in the ability of cells to generate data in an mRNA-Seq assay; such a bias would undermine the validity of this type of scRNA-Seq approach since apparent transcriptome would be defined more by how much a cell is read in the assay, rather than its true expression patterns.

The second aberrant prolife is shown in the bottom-right panel of Figure 6.11. This sample (OV04-366 $EpCAM^{DIM}$) does appear to have a genuine excess of cells. This has resulted in the total read pool being spread too thin, so UMI counts are predominantly only in double figures. With so few UMIs, it is hard to reliably determine which partitions are human cells, and which are mouse.

Another overall issue with the dataset is the degree of mouse contamination. This arguably subdivides into two discrete issues. The first is that there is a low, but extremely prevalent, degree of mouse contamination in almost all samples. This is surprising as this has not been

observed in previous flow analysis experiments. This contamination is relatively unlikely to be a result of poor gating as it is still observed in EpCAM$^+$ and CD90$^-$ subpopulations. Since these are strongly positive for EpCAM, while unstained and mouse samples are strongly negative, this would necessitate these cells having shifted across the entire axis in a repeatable fashion across 8 experiments in a manner that has not previously been seen. An alternative explanation might be low-level mis-sorting by the flow apparatus, which would explain a relatively consistent ∼10% mouse contamination across the board; however this explanation is inconsistent with the data from a pair of samples which show near-zero mouse contamination. Curiously, both are from the CD90$^-$ subpopulation. The origin of this effect is currently not clear, and no hypothesis adequately explains the observed data.

Since assignment of transcriptomic sequence to the right species is typically reliable, this low level contamination is not a significant obstacle (since the data for these mouse cells can just be discarded). The only issues are that (a) this results in a small proportion of data being lost, and (b) how this effect comes about remains unclear, which leaves unanswered questions with regard to experimental design.

The second part of this problem, is the apparent degree of mouse contamination in the EpCAM$^{\mathrm{DIM}}$ subpopulations. Unlike the majority of the samples which show low-level mouse contamination, the majority of EpCAM$^{\mathrm{DIM}}$ samples show more mouse cells than human. Given that this seems to be an EpCAM$^{\mathrm{DIM}}$-specific problem, it is likely that this is a result of poor gating during the flow sorting process.

There are several factors which may have contributed to this outcome. The EdU work did not require sorting, and so analysis of the data, and delineation of the boundaries of each subpopulation could be established without time-pressure, post-hoc, and with the benefit of a complete dataset. In this case, gates had to be applied, using a small analysis set, prior to the sample being run. In this case, things were also complicated by the use of unfamiliar software on the sorter, which made accurately judging event-densities (and thus the likely bounds of the subpopulations) much more challenging. Finally, due to logistical limitations, the scRNA-Seq was performed without any 'dry run' sort of samples in order to optimise and calibrate parameters and gates for each sample. While most of these decisions were forced by logistical concerns, the lack of preparation and optimisation proved to be a major obstacle to efficient data collection (sequencing) and analysis.

With the problems above mapped out, and pilot sequencing performed, more accurate estimates of cell number in each population could be generated. Based on this, samples could be optimally pooled in order to effeciently sequence high read numbers. From the type of plots shown in Figure 6.11, I was able to estimate usable cell numbers using assignments based on ratio of human UMIs to mouse UMIs, with some inspection and manual input for the edge-cases shown above.

The CRUK Cambridge Institute Bioinformatics Core performed a similar analysis using more in-depth methods. These analyses were performed direct from the raw data (i.e.

circumventing Cell Ranger entirely, so that all analysis processes were understood). In addition, these processes considered not just UMIs, but total read numbers for each partition. This is important because it gives an idea of the quality profile of the data, and the relative expenditure of sequencing (in terms of absolute read number) across the profile. Low quality human cells (i.e. those with low UMI counts) will not generate sufficient data to allow for meaningful analysis and must be discarded in QC steps. Consider two hypothetical samples in both of which 10% of (non-empty) partitions contain mouse cells, 60% contain low-quality human cells, and 30% contain high-quality human cells. In both cases, only 30% of these partitions will be usable for analysis, and all data for the others will be discarded. If in sample 1, 90% of sequencing reads are found in the 30% of usable cells, then very little sequencing data will be discarded. However, if in sample 2 only 10% of the total reads for this sample come from partitions containing usable cells, then 90% of sequencing data is going to get discarded. If this case, a 10-fold higher concentration of sample 2 should be added to the final pool for sequencing such that both samples will get the same number of usable reads. This hypothetical example, demonstrates the danger of this approach – that a handful of poor quality samples become a massive drain on the total pool, and can swallow up a vast proportion of the total sequencing available.

The final estimation of viable cell numbers was performed by ranking partitions (inc. 'empty' partitions) from most human reads (rather than UMIs) to least, and selecting the top 1%. Despite being a relative measure, this top-percentile resulted in fairly similar absolute quality cut-offs in all samples. The proportion of the total reads available going to this fraction of partitions could then be determined. A pooling strategy was then established by calculating the relative concentration of each sample which would need to be pooled such that each sample would be expected to give the same mean number of reads.

The resulting pooling strategy was similar to that developed from my own calculations based only on UMI count and manual interpretation; the derivation of these answers by two broadly independent methods, suggests the conclusions are robust. The primary issue highlighted by these analyses is the problem posed by samples with large proportions of mouse cells and / or low-quality human cells. In samples where a very low proportion of usable reads are going to high-quality human cells, the scaling factor required to get mean-reads-per-cell up to the sample-average is such that a small number of these samples demand the vast bulk of the available sequencing resources.

In order to mitigate this problem, two steps were taken. Firstly, two samples with extremely low quality (OV04-449 $EpCAM^{+}$ and $EpCAM^{DIM}$) were discarded entirely, on the grounds that the sequencing resources required to salvage them would be better spent on the remaining samples. Secondly, another five samples with quality noticeably lower than the rest of the samples were pooled separately. Since their scaling factors were so much higher than the rest of the samples, small margins of uncertainty could result in these samples taking a drastically smaller or larger proportion of the allocated sequencing pool than expected. As such, these samples were sequenced separately in a defined set of sequencing lanes, so the

amount of sequencing they would take could be clearly prescribed. In this case, the risk of these samples draining reads away from the main pool appeared a more pressing factor than any minor batch effects that might result from sequencing in independent lanes.

For the second round of sequencing, 8 HiSeq lanes were used. Of these, 3 lanes were allocated to the 5 low-quality samples, and 5 lanes to the remaining 21 high-quality samples.

While this way underway, the CRUK Cambridge Institute Bioinformatics Core performed some initial investigation into the (distinctly limited) data available from the first lane of sequencing. A summary of the analysis performed on aggregated data (i.e. combining data from both sorted subpopulations and the unsorted) for the first sample (OV04-072) is shown on Figure 6.12. Each of the figures shown here reveal some interesting things about the data. The first plot shows UMI counts across the population. It is immediately apparent that the unguided clustering t-SNE algorithm is to a certain degree separating cells on the basis of quality (UMI count). The large swathe of blue in the top right, and to a certain extent the bottom left corner, (marked with blue arrowheads) appear to be distinct in having very poor numbers of UMIs, and thus by definition a narrow expression spectrum. For this reason, these regions can reasonably be neglected. The same plot mapped by subpopulation of origin (top-right) provides insight into the remaining regions. The region on the left of this plot (second-from-bottom; red arrowhead) appears to contain a mix of $\text{EpCAM}^{\text{DIM}}$ cells and unsorted cells, while the major central region and bottom right (green arrowheads) contain a mix of EpCAM^{+} and unsorted. This is encouraging as it suggests that the former are $\text{EpCAM}^{\text{DIM}}$ cells and the latter EpCAM^{+} cells, with both naturally being found in unsorted sample. The graphs at the bottom show the same populations painted with their expressions of *EpCAM* and *THY1*. This appears to support these conclusions, with EpCAM being less transcribed in the population corresponding to $\text{EpCAM}^{\text{DIM}}$, but high in that corresponding to the EpCAM^{+} subpopulation. Expression of CD90 is somewhat less clear-cut (not unexpected given that expression of this marker appears proportionally lower than EpCAM), however it does appear to be less expressed in the cells corresponding to the $\text{EpCAM}^{\text{DIM}}$ subpopulation. More intriguingly, its expression appears reduced in the population in the bottom-right, despite EpCAM being strongly expressed here. This raises the tantalising potential that this population may correspond to a small CD90^{-} subpopulation. The flow data for this sample shows a thin scattering of events in the quadrant corresponding to where a CD90^{-} subpopulation would be expected. This was considered far too disparate to make it practical to sort, however this dataset suggests that it may be possible to extract this population from the unsorted sample.

These findings were very preliminary, since they were based on a single pilot lane of sequencing. Once the second round of sequencing (eight optimised lanes of data) was complete, analysis turned to this output; this sequencing was completed in the second week of September, so analysis is still very much an ongoing process. A high-level summary of the data is presented using my R-script visualisation in Figure 6.13.

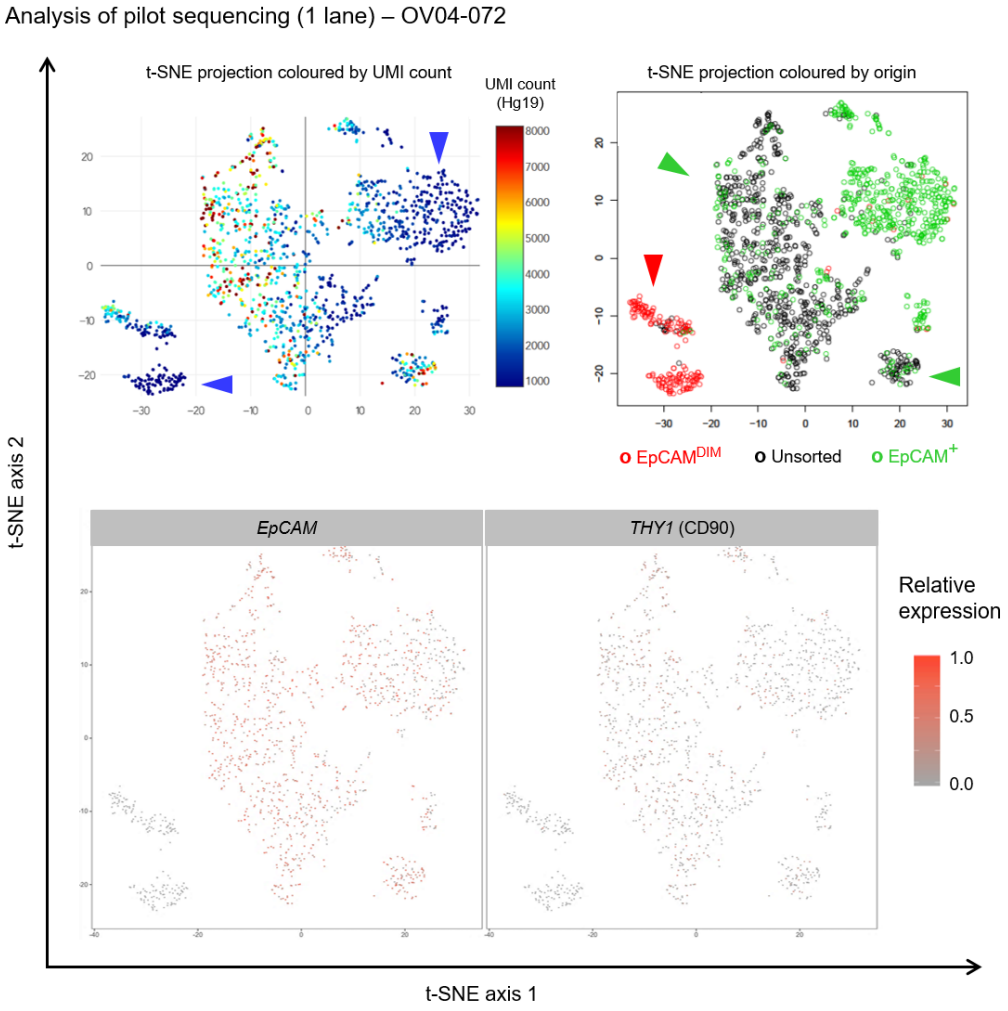


Figure 6.12: Analysis of OV04-072 (aggregated data from both subpopulation samples and the unsorted sample) from the sequencing performed in the first pilot lane only. All plots show the same unguided clustering based on the 3 aggregated datasets by t-SNE. Top-left shows t-SNE clustering, with cells painted based on total UMI-count per cell; top-right shows the same data with the data painted to indicate which subpopulation the cell in question came from; the bottom panels again show the same data with cells painted based on their expression of EpCAM and THY1 (CD90), with high expression in red and low in grey. Plots generated by CRUK Cambridge Institute Bioinformatics Core.

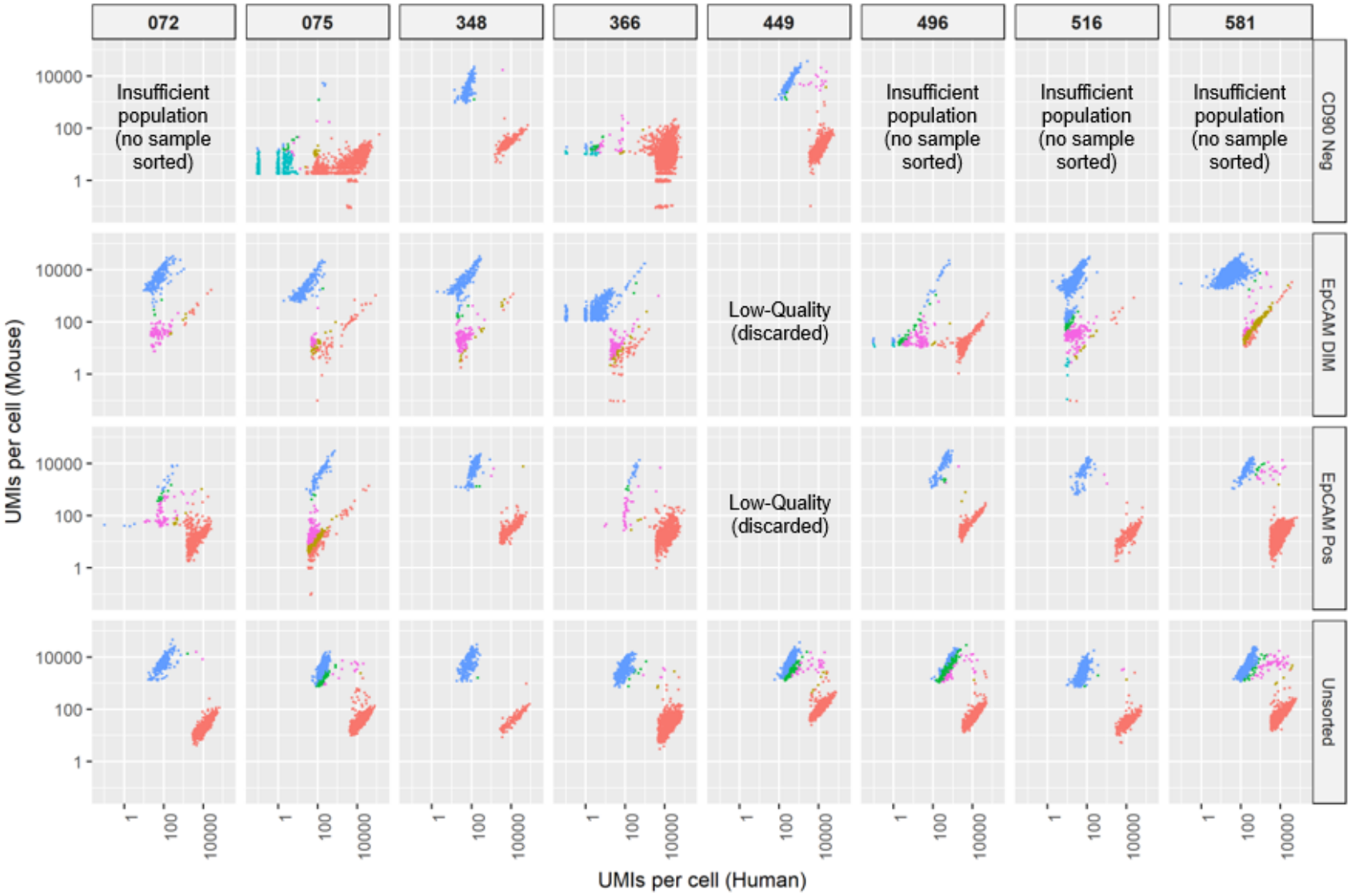


Figure 6.13: UMI count profiles (each point is a cell / partition) for second-round scRNA-Seq library sequencing. Colouring is using my own improvised calling (prior to manual inspection); cells in red are presumed human (10x more human UMIs than mouse), and cells in blue presumed mouse; gold and green are probable human and probable mouse respectively, purple is likely multiplet, and turquoise too poorly sequenced to call.

The immediate problem this reveals is clear undersequencing of most of the EpCAM^{DIM} subpopulations. This may be something inherent in their sample quality or biology, or it may be a result of suboptimal allocation of power to the two separated sequencing pools. Two of the EpCAM^{DIM} subpopulations (-075 and -496) were included in the main pool. Subsequent investigation of the sorting profile of -075 raises questions as to the validity of the 'EpCAM^{DIM}' that was sorted from this sample (a similar problem to that seen with -449, resulting in that sample being discarded). As such, the poor quality of this EpCAM^{DIM} sample is likely to be explained by poor biological quality of the sample and is therefore likely to be sample-specific. The other sample (-496) shows drastically better sequencing quality than the rest of the samples; this indicates that the low UMI counts in the other EpCAM^{DIM} populations may be because these samples were (comparatively) starved of reads.

Looking across the human cells (shown in red in the UMI plots) within the EpCAM^{DIM} populations, there appears to be an extremely wide range of UMI counts emerging from the sequencing. An example from -496 EpCAM^{DIM} is shown expanded in the top panels of Figure 6.14, plotted by UMI count and by t-SNE heatmap. As can be seen from this, there are a minority of cells getting tens-of-thousands of UMIs, while the majority of cells appear to languish in the vicinity of 5k UMI. However, this does not appear to be problem unique to the EpCAM^{DIM} samples; a similar effect is seen for many of the EpCAM^{+} samples too (see equivalent data for -496 EpCAM^{+} shown in the bottom panel). From the t-SNE plots, it is clear that this is the primary factor driving t-SNE clustering in these samples.

At this point, the natural question is whether this represents a real effect, or an artefact of the experiment. From the summary shown in Figure 6.12, it appears that the effect is broadly seen across all samples and all subpopulations – thus is not an issue with poor biology in a specific sample, or an issue with sorting.

A key question would be whether these different populations actually have different numbers of UMIs, or if there is simply a disparity in the degree of sequencing these populations are receiving. An indication of the probable answer to this is visible in Figure 6.14, simply from the shape of the plot. The UMI counts for both these samples are diagonal, indicating that as the number of human UMIs detected increases, so does the number of mouse UMIs. It is possible that this occurs due to mRNAs released from dead or dying mouse cells finding their way into partitions containing human cells, and being included in the library preparation for this partition. If the distinction between low- and high-quality cell populations was a true biological effect (i.e. the low-quality cells actually have a very low number of human UMIs), then it would be expected that their uptake of these contaminants would be random, regardless of the cell's UMI count. As such, this would suggest that the issue is one of sequencing saturation – there are too few reads, and as such the majority of UMIs in a cell are not being sequenced – different cells naturally receive a range of sequencing depth, and those which have more (human) UMIs sequenced, receive a commensurate degree of sequencing of contaminant mouse UMIs, resulting in the diagonal pattern seen. If there were really human cells with low UMI, these would expect to receive a random degree of

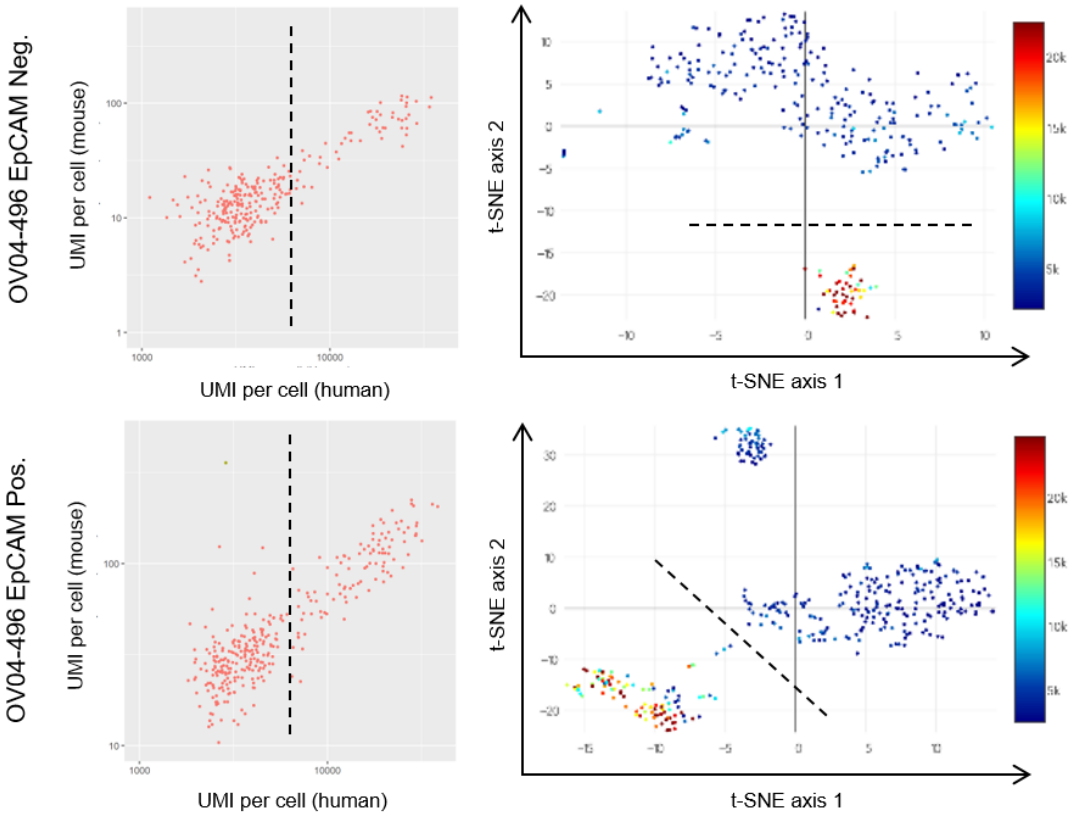


Figure 6.14: Left-hand plots show UMI count for cells from the two separate samples (top-left and bottom-left), with number of UMI mapping to the human (x-axis) and mouse (y-axis) genomes respectively. The right-hand side in each case shows the same data clustered on a t-SNE plot and painted based on total UMI count for each cell. The dotted line is added to emphasise how in both cases the samples appear to distinctly subdivide into high- and low-quality data based on UMI count, and is added to emphasise this point (exact placement is arbitrary). The left hand plots show how this distiction appears to arise as a result of a bimodal distribution in human UMI number. The right-hand plots show the impact that this bimodal distribution of human UMIs has on unguided t-SNE clustering. t-SNE plots generated by CRUK Cambridge Institute Bioinformatics Core.

mouse UMI contamination, and the plots should be horizontal. Another possibility is that this effect could occur as a result of randomly occurring PCR errors mutating either the partition barcode or the index barcode; this would result in incorrect assignment of a UMI to the wrong partition. In this case, every partition should receive an approximately similar tiny number of incorrectly assigned UMIs from other partitions and samples, which should result in a random distribution of mouse UMIs in human samples. Again, if the high- and low-quality effect was real, this should result in a random distribution of mouse UMIs, which is not observed. If it is a result of insufficient saturation, then rate of misassigned mouse UMIs should correlate with the rate of UMIs, again matching the observed data. In conclusion, this indicates that the sequencing depth, even in this second round, is insufficient, and indicates that a third round of sequencing is required as future work.

In the interim, the second round of sequencing was investigated, despite awareness of this limitation. For analysis of this data, the CRUK Cambridge Institute Bioinformatics Core generated an edited pipeline based on Cell Ranger in order to minimise issues with the program's calling algorithms. The objective was to avoid loss of reads due to misassignment to the wrong genome. To do this, reads were aligned to the combined mouse and human genomes (Hg19/Mm10), and a list of partition (i.e. cell) barcodes generated for those which were predominantly human (>90% reads align to Hg19 not Mm10). These partitions only were then selected and passed to Cell Ranger for alignment to Hg19 only. This prevents false negatives, and loss of valuable human reads/UMIs to misalignment to the mouse genome. Since there is a very low level of mouse UMIs as a result of this partition selection, the number of false positives (mouse UMIs misaligned to Hg19) is negligible.

The resulting output for two priority samples -072 and -496 is shown in Figure 6.15. These samples were prioritised due to the promising results discussed earlier, and a high UMI count in the $EpCAM^{DIM}$ subpopulation respectively. The top plots of this figure show unguided t-SNE clustering for aggregated (both sorted subpopulations plus unsorted) data for each sample. As might be expected given previous discussion, much of the dimensionality of these t-SNE plots is expended sorting low-UMI cells from high-UMI cells. In order to attempt to identify real heterogeneity in these samples, a filter was applied, discarding all cells which contained <10k UMIs. The resulting plots are shown in the bottom panel.

With the removal of the low-UMI cells, the UMI-count ceases to be a major factor in the t-SNE clustering of the remaining cells. Furthermore, this clustering appears to isolate small and distinct subpopulations of cells, suggesting discrimination of a phenotypic subgroup. In order to determine if previous results were recapitulated, similar figures were generated showing the origin ($EpCAM^{+}$, $EpCAM^{DIM}$, unsorted) of the cells. These are shown in Figure 6.16 (note the different colour-coding from the previous figure).

These plots are stark, in while the unfiltered plots largely recapitulate the previous data, application of the 10k UMI quality filter results in the loss of the previously noted pattern. There are two possible rationales for these results. Firstly, that the previous observation was an artefact, and simply a clustering of low-quality $EpCAM^{DIM}$ and unsorted cells, giving

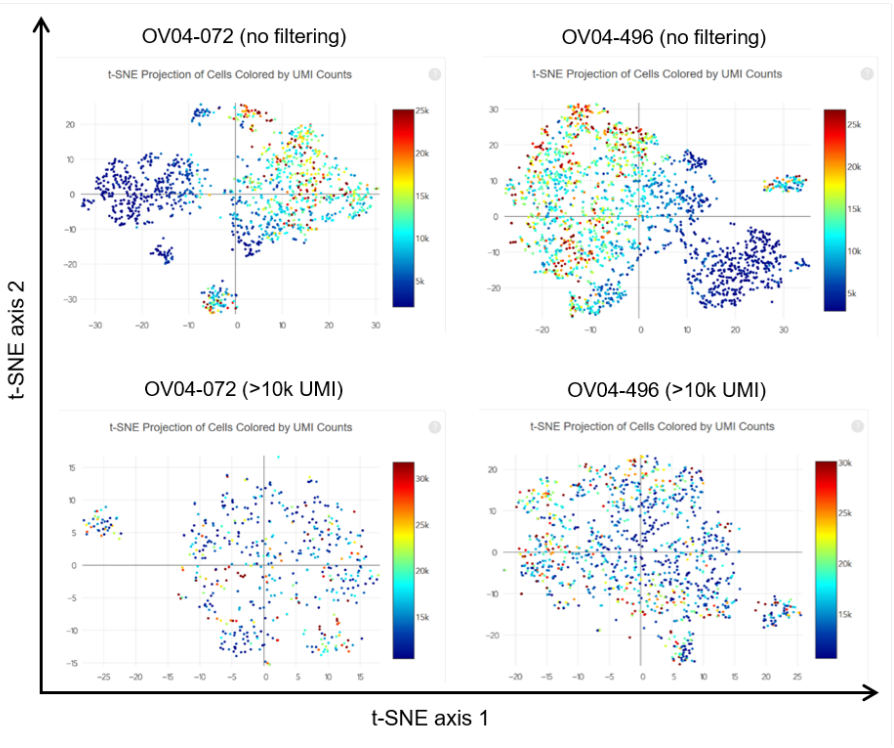


Figure 6.15: Unguided t-SNE clustering of aggregated subpopulations (i.e. EpCAM^{+}, EpCAM^{DIM}, and unsorted combined) for samples -072 and -496 from the second round of sequencing. Above and below show the same data with and without the application of a 10k filter, which discards all cells cells with <10k UMIs. The bottom two plots show how application of this filter can help mitigate the effect observed above where t-SNE clustering is heavily influenced by sequencing depth (i.e. total UMI number), rather than anything deeper within the transcriptomic profile. As can be seen in the lower plots, clustering does not appear to be heavily influenced by depth, with location on the plot and colour no longer being clearly correlated. Plots generated by CRUK Cambridge Institute Bioinformatics Core.

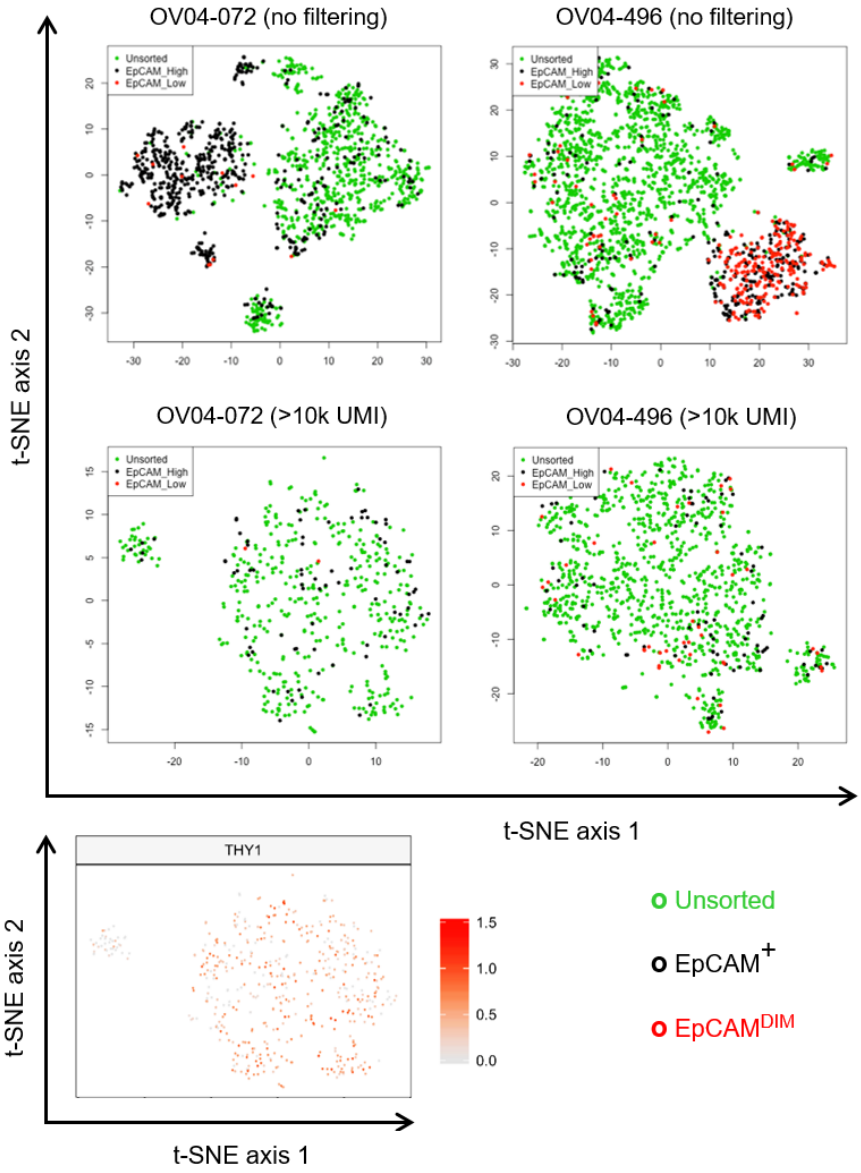


Figure 6.16: The origins (EpCAM^{+}, EpCAM^{DIM}, unsorted) of the cells shown in Figure 6.15. Note that the colour-coding varies from the previous figure. Expression for *THY1* in filtered -072 is shown in the bottom left. Plots generated by CRUK Cambridge Institute Bioinformatics Core.

the impression of a distinct phenotypic subpopulation. Alternatively, the inverse causation is true, and the loss of the previously noted pattern is due to a lack of sequencing saturation in the EpCAM^{DIM} cells, resulting in the artefactual loss of most of the data from this subpopulation.

Given the apparent lack of saturation in the samples, read depth will have a very strong impact on UMI count. Since we already know that saturation appears to be lower in the sorted EpCAM^{DIM} subpopulation, due to the unbalanced pooling, the application of an absolute 10k cut-off for cells under 10k UMI arguably results in not comparing like-for-like across multiple samples.

On a more positive note, the filtered -072 shows a small resolved population on the left of the graph. This population shows a noticeable reduction in expression of *THY1*, suggesting that it may represent the CD90^{-} subpopulation noted in the pilot dataset. A pair of similar subsets appear visible in the bottom-right of -496 although from inspection these do not appear to be noticeably depleted in *THY1*, despite the presence of a similar low-level scatter of events in the CD90^{-} region of the -496 flow profile. Objective judgement of these subsets for both samples should be reserved for after proper statistical assessment. However, if true, the ability to resolve subsets from this high-quality data would add further weight to the possibility that lack of discrimination of a EpCAM^{DIM} subpopulation results from insufficent sequencing saturation.

Ultimately, further insight into this matter is likely to require a third round of library sequencing.

6.4 Discussion

The primary objectives of this chapter of work were to:

- Expand on the understanding of established tumour proliferation developed in the previous chapter of work, by determining if subpopulations have low-level bipotency
- Further develop understanding of the EpCAM^{DIM} and CD90^{-} subpopulations by determining molecular differences between these cells and the EpCAM^{+} bulk

The most comprehensive answers to the former would be answered by lineage tracing experiments. By inducing marking in single cells, the phenotypes that arise in the resulting clonal patches can be examined. Lineage tracing provides the most complete answers because it will read-out not only when one subpopulation gives rise to another as a result of asymmetric division, but also when one type of cell gives rise to another by interconversion of existing cells. The weaknesses of this technique arise from it being a relatively indirect method of study, meaning conclusions must be extrapolated statistically; all a mixed clone

indicates is that at least one cell has changed its phenotype at some point during the study. Additionally, it is hard to definitively prove that rare events have not come about due to aggregation of two marked clones.

An alternative approach is to assess division symmetry directly. This approach would allow elucidation of the rates of asymmetric division within phenotypic subpopulations. The benefit of this appraoch is that thousands of individual independent divisions can be assessed at once, allowing a statistically sound picture of division patterns to be built up quickly. This technique also does not necessitate the time and complexity of introducing a means of lineage tracing into cells. The major downside of this technique is that it can only detect phenotype switching as a result of asymmetric division, and will miss any interconversion by existing cells; although a consequence of this is that unlike lineage tracing, the mechanism of switching will be obvious by virtue of the fact it is detected by this type of assay. This assay is necessarily dependent on the assumption that asymmetric division involves the asymmetric localisation of key proteins during the process of division. This assumption is of course not guaranteed, however previous studies have shown this to be a demonstrable phenomenon within processes of epithelial replenishment (Clayton et al., 2007), and thus investigating its relevance to the HGSOC context seems a logical enquiry.

The application of lineage tracing to the context of HGSOC PDXs is unusually complex for a number of reasons. Firstly, lineage studies are typically performed in model mice, with tracing constructs being inserted at the germline level. For PDXs this is obviously not an option, and a complex inducible construct must be inseted by viral means (or similar). This is particularly challenging in ovarian cells due to difficulty of sustaining them *in vitro*, and of transducing these cells. This adds significant technical complexity to an already time-consuming experiment, and as such it has not been possible to perform such lineage tracing within the scope of this project.

In the context of the logistical limitations of lineage tracing, the symmetry assay becomes an appealing approach to elucidating additional information about exchange between phenotypic subpopulations. While it has not been possible to complete the technical development of this method within the available timeframe, the remaining hurdles are technical in nature and theoretically minor. While it is not possible to demonstrate a proof-of-concept at this time, this approach offers the potential to generate data within the immediate term, something which lineage-tracing cannot without significant development of suitable constructs and viral vectors.

The second part of this chapter aimed to investigate the pathways underlying the different phenotypes. Single-cell mRNA sequencing allows the differential transcriptomes of the phenotypic subpopulations to be investigated, and provides the molecular insight that will allow future studies to specifically target phenotypic subpopulations.

The scRNA-Seq work described here represents another ambitious attempt to apply a cutting-edge technique in a complex and challenging context. This technique has not tra-

ditionally been applied to xenograft models, and this has resulted in a range of technical challenges which have made data collection significantly more complex. This experiment has allowed these issues to be examined, and optimisations and work-arounds to be developed for application of this technology to the xenograft context; this represents valuable method development. The inevitable consequence of this is that the efficiency of data-collection from the experiment has been reduced; however despite this, the original objectives of the experiment have largely been met, despite the need for more sequencing and analysis than anticipated.

Present indications are that current sequencing is somewhat insufficient, and a third round of sequencing would be desirable. The second round sequencing results represent a sufficient pool of data to conduct some useful analysis, particularly with regard to confirming the presence of the suspected $CD90^{-}$ population in the high-quality data. If statistical significance is observed in expression of *THY1* (and obvious artefactual origins for this distinction can be excluded), then a comparison of the gene signatures between the $CD90^{-}$ and majority of tumour cells should provide valuable information on the differential expression of this cryptic subpopulation. At present the majority of $EpCAM^{DIM}$ cells appear undersequenced, and further sequencing will likely be focused on this group of cells. Without sufficient depth on this subpopulation, there is limited value in further speculation based on limited data. However, the apparent ability to resolve the $CD90^{-}$ subpopulation is encouraging, especially as this population is less clearly discriminated by flow cytometry, compared to the $EpCAM^{DIM}$ subpopulation. This suggests that once the data for the $EpCAM^{DIM}$ cells is brought up to the level of the $EpCAM^{+}$ and $CD90^{-}$ subpopulations, it is likely to be possible to resolve it by differential expression.

Chapter 7

Discussion

7.1 Synopsis

The aim of this project was to gain a deeper understanding of phenotypic heterogeneity within HGSOC, determine the proliferation patterns within the disease, and consider the implications of these with regards to treatment.

Previous studies in the field have identified a number of phenotypes which apparently enrich tumour-initiating cells. Many of these groups report this as evidence of a 'cancer stem cell' population within HGSOC. While the CSC theory of HGSOC enjoys strong support, there is a distinct lack on consensus over the identity of the subpopulation in question, with conflicting data from different groups, and a shortage of markers which can uniquely identify the tumour-initiating fraction, rather than simply demonstrate enrichment (Zhang et al., 2008) (Alvero et al., 2009) (Szotek et al., 2006) (Wei et al., 2010) (Curley et al., 2009) (Kryczek et al., 2011) (Silva et al., 2011) (Stewart et al., 2011).

From previous experiments in the Brenton Laboratory we have evidence for the following:

- EpCAM$^-$ cells are uniquely able to regrow tumours in *in vitro* and *in vivo* assays
- EpCAM$^-$ cells are enriched by chemotherapy

Almost all previous work has been conducted in the form of tumour-initiating studies, in which cells are subdivided based on phenotypic traits, and their growth potential assayed in *in vitro* and *in vivo*. Tumour-initiating studies presume comparatively strong growth (or indeed all observable growth) from a subpopulation is an indication of proliferative capacity. However, this assay is inherently also measuring resilience. Cells which do not survive the transfer from host to model environment will not show indications of growth. Equally, cells' interactions (and adhesion) with their environment are also important, and

may determine their ability to initiate proliferation. Thus the assay is also measuring a cell's capacity and/or requirement for adhesion relative to the ability of the model to replicate this. The importance of cells' external interactions is highlighted by the previous work in the Laboratory showing the inability of isolated HGSOC TICs to engraft, even in large numbers, while mixed populations grow easily. Thus while these experiments provide interesting information about cells' comparative abilities to regrow tumours (which may potentially be informative as to the processes occurring in tumour metastasis and recurrence from minimal residual disease), they do not necessarily provide any reliable information about the growth processes occurring within an *established* HGSOC tumour.

The broad goal of this project has been to investigate the limitations of the tumour-initiating approach, by developing and studying xenograft models to see if this approach would generate different results. By experimental interrogation of proliferation within these models, the dynamics of proliferation can be studied in an *in situ* context.

During this project, 12 PDXs were successfully grown, passaged, validated (by histology, TAm-Seq, and sWGS), and used for subsequent experimentation. Additionally, over the course of this process, selection and engraftment practices were refined to the point where the majority of samples being engrafted were successfully implanting; indeed by the end of the project, several samples were stored rather than being passaged further, as there was little scope for the use of further biological replicates. Thus not only is a bank of samples available for future work, but there now exists the means to generate more, including from specific samples of interest. This option is currently being utilised, with another project in the Laboratory generating PDXs from untreated primary patient samples.

With the improved engraftment methodology, solid tumour samples have proved to engraft efficiently. However, due to the delay in developing these improvements, ascites (which proved easier to engraft) ended up making up a good proportion of the PDX samples. These cells represent both advanced disease, and disease that has entered the metastatic phase. This explains the improved engraftment rate, but also raises some question over how truly representative these samples are of HGSOC as a whole (and particularly of earlier stage disease). Thus, while these samples represent important models of late-stage disease, it is important that any results are demonstrated to be valid in samples of both solid tumour and ascites origin before being said to be true of HGSOC as a whole.

The use of purified ascites spheroids was also investigated as a way of improving the efficiency of model generation. Surprisingly, these proved to be less efficient than equivalent whole ascites. This raises questions as to whether the assumptions about spheroids (namely that the presence of clonal growth indicates strong proliferative capacity and/or the presence of a cancer stem cell at their core) are entirely valid, as these cells do not appear to be enriched for tumour-initiating efficiency. Potentially this may indicate these cells are actually less independently resilient and are dependent on the cells which surround them (which may not survive transplantation). On the other hand, it may simply be an artefact of the experiment. Potentially the filtration process may impose stress on the cells; alternatively, it may be that

there are important interactions between spheroids and free ascites which is lost when the two are separated, leading to reduced viability. Mechanical digestion was also investigated as a method for implantation. This proved variably efficient, and notably produced very low viability in the cells being transplanted. As such, it may be a valuable tool for passaging tumours in a pragmatic manner; however it is unclear whether this has any impact on how representative or heterogeneous the resulting tumours are (as a result of the very low viability of the implanted material). Due to these concerns, this approach was not widely applied in this project. In hindsight, given the high cell attrition rates the method was later determined to have, this was a good decision. While the added resource and logistical burden from implantation is significant, where this is not prohibitive it is strongly preferable to use this approach in studies of heterogeneity, given the potential scope of the impact this attrition on complexity.

When growth within xenografts was assessed by nucleoside incorporation experiments, marking was observed to be stochastic. When analysed by flow cytometry using the markers EpCAM and CD90, these xenografts subdivide into two or three subpopulations. Where only $EpCAM^{DIM}$ and $EpCAM^{+}$ are observed, the $EpCAM^{DIM}$ subpopulation is shown to proliferate more slowly in both short-term and long-term experiments, with both subpopulations retaining stochastic division profiles. That the $EpCAM^{DIM}$ cells are marked at a slower rate in both assays indicates that this subpopulation is both dividing more slowly, and being replenished more slowly. The similar values for median division time that are calculated from both assays (for both subpopulations respectively) also strongly suggest that these subpopulations are self-supporting, since it is unclear why they would mirror each other reliably under other circumstances. Analysis of the EdU uptake data, in conjunction with the growth rates of xenografts (estimated from the size of tumours over time), reveal that the proliferation rates of the $EpCAM^{DIM}$ and $EpCAM^{+}$ populations are not compatible with a cancer stem cell model. These data demonstrate that all significant proliferation within these subpopulations appears to arise within the subpopulations themselves.

It is not possible to conclude anything from these data about any possible low-level transition of cells between the subpopulations. This information could be important in considering the dynamics of tumour subpopulations where one subpopulation is significantly depleted, as appears to occur during tumour-initiation and chemotherapy. However, given that it can be concluded that such dynamics do not play a significant role in *in situ* growth, it would seem disingenuous to describe these cells as 'cancer stem cells', rather than noting that low level phenotypic interconversion may be possible.

In order to clarify this matter, lineage tracing methods were investigated. Given the extensive technical challenges involved in this process, it did not prove possible to develop this methodology within the timeframe of this project. The largest hurdle to this process was the development of a suitable lineage tracing construct for viral transduction into xenograft cells. Cells successfully transduced with a zsGreen control vector were implanted into a mouse, and in the resulting tumour, the majority of the cells were green; this provides a

proof-of-concept for the relative ease of the latter stages of this work. In this project the main challenge was that all of the constructs that were tested did not have the expected properties. More generally, the technical limitation is finding a viable balance between a construct large enough to contain all of the necessary genomic components, such that permanent lineage marking can occur in an inducible fashion, while also keeping the construct small enough to allow efficient transduction in ovarian cells. Continued development work may be able to overcome this problem, however such optimisation is a very uncertain process, and it is not known how protracted such method-development might be (in addition to the lengthy nature of the studies themselves). Additionally, there is no guarantee that a successful solution will be obtained with current technology.

In an effort to obtain partial information about possible interconversion processes without the need for an indefinite period of method development, the practicality of studying division symmetry by IF or IHC was investigated. By marking pairs of recently divided cells, the ability of cells to undergo asymmetric division can be interrogated. Unfortunately, the technical optimisation for this methodology has also been complicated, and has not been completed in the short timeframe available. However, the challenges remaining are significantly less complicated than those for lineage tracing, and these challenges are solvable within a timeframe of months rather than years. This option provides more limited information (in that it can only detect asymmetric division rather than interconversion from one phenotype to another), but offers a much more highly powered assessment of the dynamics of division itself.

In order to investigate xenograft subpopulations at the molecular level, scRNA-Seq was performed. Limited time has allowed for only a preliminary analysis at this stage, but there are promising indications of an ability to resolve $CD90^-$ cells from the $EpCAM^+$ bulk. Analysis of differential expression patterns can now be performed on these data sets. At present there appears to be insufficient power in the $EpCAM^{DIM}$ subpopulations, to perform comparison with the bulk, but another round of sequencing should raise the data quality of these cells to a point where they are equivalent to the $CD90^-$ subpopulation, and can be subject to analysis of differential expression in a similar fashion.

7.2 Conclusions

7.2.1 Proliferation

The key issue for the clinical treatment of HGSOC is relapse and resistance. Patients typically receive chemotherapy and respond well, but subsequently suffer disease relapse and increasing resistance to chemotherapeutics. This has naturally led to significant interest in the biology of resistance, and how and why there is a trend for treatment to become less effective over time.

The CSC model (a model for which there is strong evidence in other tumour types (Barker et al., 2007) (Chen et al., 2012)) offers a tempting and elegant solution to this problem. Chemotherapy could kill bulk and TA cells, leaving a resistant CSC population which could then regenerate the niche. A wealth of previous work (noted in section 7.1) has attempted to identify CSCs based on numerous cell-surface markers, and has demonstrated enriched TIC capacity in these subpopulations.

There are a number of key flaws in such an argument however, including evidence of extensive proliferation in bulk tumour (Chen et al., 2017), that TIC capacity does not translate into *in situ* proliferative capacity, and that a CSC model arguably explains accumulation of resistance with treatment less well than alternative models (since a resistant CSC should always regenerate the same sensitive progeny, while in a more stochastic model, treatment would act as a stronger selection pressure for resistant subclones).

The results of this work demonstrate that proliferation within HGSOC (within the limits of xenograft models) does not arise from a small population of cells, and instead arises from the population as a whole. This is fundamentally incompatible with a CSC model, and strongly suggests proliferation is stochastic in nature.

7.2.2 Chemoresistance

These results go a long way to explaining the varied and in some cases apparently conflicting nature of the stem cell markers proposed in the existing literature, suggesting that in identifying markers enriching for TIC capacity, what these authors have actually done is found markers which enrich for cells which retain a greater degree of their stochastic proliferative ability in a stressful transplantation assay. Since tumour samples may differ depending on origin, it is possible that a marker that enriches for such cells in one study's samples will deplete then in another, providing a plausible explanation for these contradictory results. Effectively these studies appear to have isolated cells with high tolerance for stress.

However, considering the ultimate clinical context of this work – relapse and resistance – this question is far from an invalid one. Given that we now know that proliferation is not limited to any particular part of the tumour, it would seem reasonable to conjecture that the development of resistance in the clinical context results from selective elimination of all but the most chemoresistant cells. This may occur by selection of a naturally variable cell population (with these cells giving rise to a tumour composed of clones with a slightly higher baseline of resistance). Alternatively, it may arise as a result of enrichment of naturally occurring distinct phenotypic subpopulations with intrinsic resistance to chemotherapy.

Evidence from this study suggests there is an $EpCAM^{DIM}$ subpopulation which divides more slowly than the $EpCAM^{+}$ bulk. Since cisplatin / carboplatin therapy targets dividing cells, it would be reasonable to expect that these cells might be less sensitive to treatment, and thus be enriched in treated populations. Pilot work (not shown) appeared to show that in

mice undergoing cisplatin treatment, the fraction of each xenograft tumour subpopulation marked with EdU remained at near-zero over the first few days of tracing, in stark contrast to the untreated samples. From this, it might be conjectured that cisplatin is disrupting DNA replication in cells attempting to proliferate, either preventing significant EdU uptake, or killing the cells outright, meaning they take up DAPI and are eliminated from the analysis. As such, attrition would be expected from cells of all subpopulations, but with greater loss from those proliferating faster. This would result in a proportional enrichment of the slowest dividing subpopulations. As cell number reduces, the proportion of $EpCAM^{DIM}$ cells would inexorably increase, ultimately resulting in $EpCAM^{DIM}$ eventually becoming the dominant population. Once chemotherapy is stopped, this process would reverse, and $EpCAM^{+}$ would begin to outgrow the $EpCAM^{DIM}$ subpopulation. Both of these outcomes are typical of what is observed in clinical samples.

7.2.3 Previous work

Further conclusions regarding the effects of chemotherapy versus EpCAM status are hard to rigorously draw, since the majority of data comes from previous work by Siru who studied somewhat different models of these subpopulations. At a high-level, these findings appear broadly consistent with data from Siru's work which suggest that $EpCAM^{+}$ cells are significantly depleted by chemotherapy in xenograft models; however her work in xenografts was very preliminary, and it is hard to draw solid conclusions from the limited data available, particularly as there is little validation of the xenograft populations. Siru's preliminary xenograft work is interesting in that despite proposing an $EpCAM^{-}/PDPN^{+}$ CSC population on the basis of TIC capacity, her xenografts do not show evidence of this population. She does observe $EpCAM^{+}$ cells with some degree of PDPN positivity, but not her proposed TIC subpopulation. These observations are very similar to my own, and are not too dissimilar to what is seen when PDPN is replaced with CD90 as an alternative marker.

Potentially, this may arise as a result of podoplanin expression being abolished in xenografts due to lack of proper human signalling from mouse stromal cells. However, I would tentatively advance an alternative hypothesis in which the bulk of this $EpCAM^{-}/PDPN^{+}$ seen in many primary samples represents cancer associated fibroblasts (CAFs) with low level contamination with tumour cells. Human stroma would be $EpCAM^{-}$, and cancer associated stroma is known to contain podoplanin positive cells. We know that tumour cells suffer intense stress as a result of transplantation, resulting in severe attrition and loss of ability to proliferate. The presence of $PDPN^{+}$ CAFs may provide a niche for any tumour cells, providing an environment in which cells are able to survive and proliferate in a more typical fashion. In support of this, there is extensive literature suggesting that podoplanin-enriched CAFs can promote tumour progression and poor prognosis (Schoppmann et al., 2012), promote EMT-like properties in tumour cells (Yamanashi et al., 2009) (Kawase et al., 2008), and (crucially) can promote tumour-initiation *in vivo* (Hoshino et al., 2011). Since we now know that any and all HGSOC cells are (under *in situ* circumstances at least) capable of

proliferation, the necessary number of contaminating tumour cells does not need to be very high to allow a tumour to be regenerated, assuming those cells are able to survive the transplantation process. This hypothesis might initially appear inconsistent with Siru's zsGreen labelling experiment (described in Figure 1.2), until it is considered that the cells which were not virally labelled were also incubated for >12 h *in vitro* while the EpCAM$^-$/PDPN$^+$ subpopulation was transduced with virus. If unsupported tumour cells are not able to remain viable during this transplantation (due to the absence of PDPN$^+$ CAFs), then it is entirely reasonable that the few tumour cells in the labelled population (surrounded by PDPN$^+$ CAFs) are able to subsequently outgrow them. The final (and arguably most convincing) piece of evidence for this hypothesis is that when Siru conducted digital PCR to validate the presence of tumour in the proposed CSC population, the percentage of mutant alleles was observed to be only ∼1%. In her work this was assumed to support her hypothesis because the small fraction of cells positive for p53 mutation in the digital PCR broadly correlated with the fraction of cells with TIC capacity *in vitro*, however with the work laid out in this project providing not just an absence of evidence for CSCs but evidence of their absence, it becomes reasonable to question whether the properties demonstrated by these TIC cells are intrinsic to their phenotype, or simply conferred by their environment and surrounding cells.

The single observation that does not fit with the above is that this EpCAM$^-$/PDPN$^+$ subpopulation does not appear to proliferate in isolation *in vivo*. If it is indeed composed of CAFs along with a small number of tumour cells, it would be expected to proliferate in isolation. However, Siru shows that when isolated, this population does not grow, despite doing so when reunited with its (apparently non-proliferating brethren). This piece of evidence is hard to reconcile with the rest. One potentially plausible hypothesis is that this may occur as a result of contaminating tumour cells in the EpCAM$^-$/PDPN$^+$ subpopulation needing some kind of interaction with a sufficient density of other tumour cells, which they may still be able to receive even from the stressed and non-proliferating tumour cells grown *in vivo*.

There are still a number of questions lacking clear-cut answers over the composition of HGSOC, and the issues discussed above also highlight the importance of further research into the influences of the niche environment. Further work in this area is clearly needed, both in patient samples and in xenografts, and particularly over how this composition is influenced by chemotherapy. The xenograft models developed as part of this project are all from patients pre-treated with chemotherapeutics, and therefore are not suitable for studying changes as a result of chemotherapy, since many important changes may have already occurred. Further studies require PDX models generated from primary untreated patient samples. These are rare, valuable, and likely represent less advanced disease, which would be assumed to be harder to engraft. This highlights the value of the method-development work undertaken in this project, since the procedures established can now be used to develop PDX models in an efficient fashion without wasting this crucial clinical material. This will be discussed further in section 7.3.

7.2.4 Phenotypic nature

This project has identified two, probably three, distinct phenotypic subpopulations within HGSOC tumour xenografts, and demonstrated apparent distinct proliferation rates arising from these. Proliferation studies answer many questions about growth rates, and dynamics between these subpopulations (or as in this case, lack thereof). What such studies do not do is provide any detail as to the nature of these cells, or their underlying biology. The loss of EpCAM in the $EpCAM^{DIM}$ subpopulation implies loss of epithelial nature, raising questions as to whether these cell are EMT-like. Since many of the clinical samples used in this project were from ascites and / or from patients with metastatic disease, some EMT-like capacity would be expected in order for these tumours to spread. On the other hand, this division appears to also be seen even in patients with only localised disease, suggesting that the $EpCAM^{+}$ / $EpCAM^{DIM}$ distinction may be something more fundamental than an EMT process. Furthermore, these $EpCAM^{DIM}$ cells appear to show downregulation in many markers, not just epithelial ones (e.g. CD90, CD10, and podoplanin); this would seem to suggest they are undergoing something more akin to de-differentiation.

Single cell mRNA sequencing was performed in order to understand these cells at a transcriptomic level, and elucidate the mechanistic differences in these populations. Due to the challenges associated with the analysis of these data, and the need for additional library sequencing, this part of the project remains a work-in-progress.

When completed however, unpicking the mechanistic detail of these cells will allow the nature of the $EpCAM^{DIM}$ population to be understood more clearly. Investigating whether genes associated with active chemoresistance are upregulated in these cells will also allow determination of whether these cells are chemoresistant solely as a result of their reduced proliferation, or if this feature is conferred by distinct mechanisms.

Over the longer term (and in the context of further studies into the effects of chemotherapy), if these $EpCAM^{DIM}$ cells are determined to be the source of disease relapse, then the focus will inevitably turn to how best to target them. It is in this context that mapping the defferential transcriptomes of the phenotypic subpopulations will be extremely valuable, as it will provide a foothold for future drug-development work to develop a means to specifically target these chemoresistant cells. If these resistant cells can be pre-emptively ablated, then follow-up carboplatin chemotherapy can be used to efficiently eliminate the bulk $EpCAM^{+}$ disease.

Furthermore, single cell mRNA study should hopefully provide additional information about the $CD90^{-}$ subpopulation observed in some samples. The nature of this subpopulation is cryptic, and is discussed in limited detail because at present there is insufficent information to draw rigorous conclusions. However, a number of observed features may offer some clues as to its origin. It is observed in a relatively small number of experiments; when observed it appears to have a slightly slower proliferation rate than the $EpCAM^{+}$ bulk, but when it is present, the distinction between $EpCAM^{+}$ and $EpCAM^{DIM}$ is observed to be marginal.

It is observed almost exclusively in xenograft PDX from ascites, rather than in those from solid tumour origin. This would appear to indicate a possible link with advanced disease (or possibly that it is simply a feature of ascites cells). However, the subpopulation is observed in several PDXs derived from previous passages in which it was not seen. It is not clear if this is as a result of the subpopulation appearing only under certain circumstances, or if it is always present, but normally at sufficiently low proportion as to not be noticeable as a distinct fraction. Another possibility is that this population is an artefact which arises as a result of multiple passages. It is a permanent feature of cell lines such as PEO-4, which would be consistent with the latter hypothesis, however as these lines were developed from ascites cells this is far from conclusive. Therefore, analysis of the transcriptomic data from these cells will hopefully provide detail as to what distinguishes them, and provide plausible hypotheses which can be subsequently tested empirically.

7.2.5 Cell fate

The data from this project indicate that the rates of proliferation from the $EpCAM^{+}$ and $EpCAM^{DIM}$ subpopulations match the rates of entry of new cells into their respective subpopulations, indicating that these subpopulations are self-supporting. However, additional evidence from the individual cell level is required in order to confirm whether or not there is a low-level of exchange between these populations, either via asymmetric division or interconversion of existing cells. While Siru's work would appear to suggest that cells with low levels of EpCAM can give rise to heterogeneous tumours under the conditions of tumour initiation and wound-healing-like response, it is not clear if low-level exchange is a feature of *in situ* tumour. Perhaps more importantly, if the inverse can occur, this will have implications for any future work to target the $EpCAM^{DIM}$ population should it emerge that $EpCAM^{+}$ cells can give rise to $EpCAM^{DIM}$ cells. In this context, the findings of Janzen *et al.* (Janzen et al., 2015) are particularly interesting, as they report that labelling of the $CA125^{+}$ cell fraction (which, based on its similar properties, I believe may represent the same subpopulation as $EpCAM^{+}$), gave rise to tumours in which only these cells were marked (which may indicate that $EpCAM^{+}$ cannot gain bipotence, even in the context of a wound-healing response). However, it is worth noting some degree of reservation about the conclusions expressed by the authors of this paper. While this work very clearly demonstrates that $CA125^{+}$ cells broadly give rise to their own phenotype, there is sufficient noise in the experiment that it is questionable whether the absence of $CA125^{-}$ can be comprehensively discounted, given that 2.5% of events are still found within the $CA125^{+}$ cytometry gate.

Implantation studies provide a very distorted way to assess fate decision in *in situ* tumour, therefore further investigation of this line of enquiry would require the development of a method along similar lines to the inducible lineage-tracing, or symmetry assays described in this project. Unfortunately, at present neither of these methods is ready for experimental use.

7.3 Future work

In the immediate-term, the priority for further work will be further analysis of the single-cell mRNA sequencing data, and a likely third round of library sequencing. The main focuses of this work will be:

- Deeper sequencing of the $EpCAM^{DIM}$ population in order to extract a 'transcriptional signature' for this subpopulation
- To use information available from the sorted $EpCAM^{DIM}$ cells to identify $EpCAM^{DIM}$ cells within the unsorted samples (this allows recovery of data for the $EpCAM^{DIM}$ population from unsorted samples where the quality appears higher, and may also allow recovery of $EpCAM^{DIM}$ cells in the case of samples where the sorted population is not usable
- To use the several sets of very high-quality data from paired $CD90^{-}$ and unsorted samples to elucidate the nature of this subpopulation; however it is possible that this work may require a better working understanding of the $EpCAM^{DIM}$ population in order that this component of the unsorted can be properly neglected

A number of promising technical options exist which have the potential to provide a working symmetry-assay (for determining if phenotypic subpopulations can proliferate to give rise to cells of other subpopulations). These represent another avenue for collecting valuable data within the immediate term.

Over the longer term, an outstanding priority is to develop a more detailed understanding of how chemotherapy affects phenotypic complexity within HGSOC. In order to study this under laboratory conditions, pre-treatment models are needed. The need for phenotypic complexity requires an *in vivo* model, and the nature of the current range of PDXs (derived from tumours typically repeatedly subject to chemotherapy) adds too much uncertainty for practical experimentation. Therefore, untreated PDX models are required to allow clear comparison of both pre- and post-treatment samples and for treatment studies with drug- and control-treatment arms.

This work is already underway in a separate project within the Brenton Lab, utilising the optimised methods developed during this work. It is for this reason that this important question has been largely omitted from the scope of this project. By understanding the effects of chemotherapy on the phenotypic balance of the disease, the role of the phenotypic subpopulations in treatment and relapse can be investigated. Where work progresses from this point will depend heavily on the role of the $EpCAM^{DIM}$ population in this process.

If the stress of chemotherapy closely mimics the presumed processes occurring in xeno-engraftment, then it may well be that $EpCAM^{DIM}$ cells play a critical role in the regrowth of disease after chemotherapy. In this case, the objective of these studies must become to iden-

tify promising features of this subpopulation which could be utilised in drug-development, so that this subpopulation can be depleted prior to chemotherapy. The transcriptomic analysis resulting from this study will provide a strong foundation for such work. Alternatively, chemotherapy may not result in a wound-healing-type response; it may simply be that chemotherapy kills cells stochastically as they enter the cell-cycle, depleting the $EpCAM^{+}$ subpopulation faster than the $EpCAM^{DIM}$ to give the observed enrichment of $EpCAM^{DIM}$ cells, and that the $EpCAM^{+}$ bulk regrows from a small number of surviving $EpCAM^{+}$ cells. In this case, translational options are more limited, and research must focus on therapeutic options relevant to this new understanding of the disease. Examples of such avenues would be looking for ways to viably extend the window over which chemotherapy is given, such that a larger proportion of cells have the chance to enter cell cycle and be eliminated. Such an approach would of course exacerbate the problems with selection for cells with increased resistance, and as such the main mechanisms behind this would likely be a useful avenue of enquiry. Again, a detailed assessment of the scRNA-Seq transcriptomic data will provide a solid foundation for identifying targets which might allow pharmacological targeting of key mechanisms, in the manner of Janzen *et al.* and their work targeting cIAP with birinapant (Janzen et al., 2015).

7.4 Valedictory remarks

The primary achievement of this project has been the demonstration that proliferation in HGSOC arises from the tumour as a whole, not a specific subset of cells as per the cancer stem cell model proposed in a range of previous publications. In addition, a distinct $EpCAM^{DIM}$ subpopulation of slow-cycling cells has been identified, which appears similar to the subpopulation demonstrated to have tumour-initiating capacity in previous work. The slow-cycling nature of these cells is consistent with previous data which suggest they are chemoresistant, and may have a role in disease recurrence. Furthermore, an additional $CD90^{-}$ subpopulation has been identified in some tumours, which plays an unknown role in phenotypic heterogeneity. This project has also optimised processes for the development of HGSOC PDX models, and provided a foundation for future work investigating the effects of chemotherapy on phenotypic heterogeneity; this work which is now underway. Finally this project has collected a wealth of transcriptomic data on the different subpopulations, the ongoing analysis of which will provide further detail on the transcriptional mechanisms of these distinct phenotypic compartments.

Bibliography

Ahmed, A. A., Etemadmoghadam, D., Temple, J., Lynch, A. G., Riad, M., Sharma, R., Stewart, C., Fereday, S., Caldas, C., deFazio, A., Bowtell, D. & Brenton, J. D. (2010). The Journal of Pathology 221, 49–56.

Alvero, A. B., Chen, R., Fu, H.-H., Montagna, M., Schwartz, P. E., Rutherford, T., Silasi, D.-A., Steffensen, K. D., Waldstrom, M., Visintin, I. & Mor, G. (2009). Cell Cycle 8, 158–166.

Ang, J. E., Gourley, C., Powell, C. B., High, H., Shapira-Frommer, R., Castonguay, V., Greve, J. D., Atkinson, T., Yap, T. A., Sandhu, S., Banerjee, S., Chen, L.-M., Friedlander, M. L., Kaufman, B., Oza, A. M., Matulonis, U., Barber, L. J., Kozarewa, I., Fenwick, K., Assiotis, I., Campbell, J., Chen, L., de Bono, J. S., Gore, M. E., Lord, C. J., Ashworth, A. & Kaye, S. B. (2013). Clinical Cancer Research 19, 5485–5493.

Ashkenazi, A. (1998). Science 281, 1305–1308.

Barker, N., Ridgway, R. A., van Es, J. H., van de Wetering, M., Begthel, H., van den Born, M., Danenberg, E., Clarke, A. R., Sansom, O. J. & Clevers, H. (2009). Nature 457, 608–611.

Barker, N., van Es, J. H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P. J. & Clevers, H. (2007). Nature 449, 1003–1007.

Bell, D., Berchuck, A., Birrer, M., Chien, J., Cramer, D. W., Dao, F., Dhir, R., DiSaia, P., Gabra, H., Glenn, P., Godwin, A. K., Gross, J., Hartmann, L., Huang, M., Huntsman, D. G., Iacocca, M., Imielinski, M., Kalloger, S., Karlan, B. Y., Levine, D. A., Mills, G. B., Morrison, C., Mutch, D., Olvera, N., Orsulic, S., Park, K., Petrelli, N., Rabeno, B., Rader, J. S., Sikic, B. I., Smith-McCune, K., Sood, A. K., Bowtell, D., Penny, R., Testa, J. R., Chang, K., Dinh, H. H., Drummond, J. A., Fowler, G., Gunaratne, P., Hawes, A. C., Kovar, C. L., Lewis, L. R., Morgan, M. B., Newsham, I. F., Santibanez, J., Reid, J. G., Trevino, L. R., Wu, Y.-Q., Wang, M., Muzny, D. M., Wheeler, D. A., Gibbs, R. A., Getz, G., Lawrence, M. S., Cibulskis, K., Sivachenko, A. Y., Sougnez, C., Voet, D., Wilkinson, J., Bloom, T., Ardlie, K., Fennell, T., Baldwin, J., Gabriel, S., Lander, E. S., Ding, L., Fulton, R. S., Koboldt, D. C., McLellan, M. D., Wylie, T., Walker, J., O'Laughlin, M.,

Dooling, D. J., Fulton, L., Abbott, R., Dees, N. D., Zhang, Q., Kandoth, C., Wendl, M., Schierding, W., Shen, D., Harris, C. C., Schmidt, H., Kalicki, J., Delehaunty, K. D., Fronick, C. C., Demeter, R., Cook, L., Wallis, J. W., Lin, L., Magrini, V. J., Hodges, J. S., Eldred, J. M., Smith, S. M., Pohl, C. S., Vandin, F., Raphael, B. J., Weinstock, G. M., Mardis, E. R., Wilson, R. K., Meyerson, M., Winckler, W., Getz, G., Verhaak, R. G. W., Carter, S. L., Mermel, C. H., Saksena, G., Nguyen, H., Onofrio, R. C., Lawrence, M. S., Hubbard, D., Gupta, S., Crenshaw, A., Ramos, A. H., Ardlie, K., Chin, L., Protopopov, A., Zhang, J., Kim, T. M., Perna, I., Xiao, Y., Zhang, H., Ren, G., Sathiamoorthy, N., Park, R. W., Lee, E., Park, P. J., Kucherlapati, R., Absher, D. M., Waite, L., Sherlock, G., Brooks, J. D., Li, J. Z., Xu, J., Myers, R. M., Laird, P. W., Cope, L., Herman, J. G., Shen, H., Weisenberger, D. J., Noushmehr, H., Pan, F., Jr, T. T., Berman, B. P., Berg, D. J. V. D., Buckley, J., Baylin, S. B., Spellman, P. T., Purdom, E., Neuvial, P., Bengtsson, H., Jakkula, L. R., Durinck, S., Han, J., Dorton, S., Marr, H., Choi, Y. G., Wang, V., Wang, N. J., Ngai, J., Conboy, J. G., Parvin, B., Feiler, H. S., Speed, T. P., Gray, J. W., Levine, D. A., Socci, N. D., Liang, Y., Taylor, B. S., Schultz, N., Borsu, L., Lash, A. E., Brennan, C., Viale, A., Sander, C., Ladanyi, M., Hoadley, K. A., Meng, S., Du, Y., Shi, Y., Li, L., Turman, Y. J., Zang, D., Helms, E. B., Balu, S., Zhou, X., Wu, J., Topal, M. D., Hayes, D. N., Perou, C. M., Getz, G., Voet, D., Saksena, G., Zhang, J., Zhang, H., Wu, C. J., Shukla, S., Cibulskis, K., Lawrence, M. S., Sivachenko, A., Jing, R., Park, R. W., Liu, Y., Park, P. J., Noble, M., Chin, L., Carter, H., Kim, D., Karchin, R., Spellman, P. T., Purdom, E., Neuvial, P., Bengtsson, H., Durinck, S., Han, J., Korkola, J. E., Heiser, L. M., Cho, R. J., Hu, Z., Parvin, B., Speed, T. P., Gray, J. W., Schultz, N., Cerami, E., Taylor, B. S., Olshen, A., Reva, B., Antipin, Y., Shen, R., Mankoo, P., Sheridan, R., Ciriello, G., Chang, W. K., Bernanke, J. A., Borsu, L., Levine, D. A., Ladanyi, M., Sander, C., Haussler, D., Benz, C. C., Stuart, J. M., Benz, S. C., Sanborn, J. Z., Vaske, C. J., Zhu, J., Szeto, C., Scott, G. K., Yau, C., Hoadley, K. A., Du, Y., Balu, S., Hayes, D. N., Perou, C. M., Wilkerson, M. D., Zhang, N., Akbani, R., Baggerly, K. A., Yung, W. K., Mills, G. B., Weinstein, J. N., Penny, R., Shelton, T., Grimm, D., Hatfield, M., Morris, S., Yena, P., Rhodes, P., Sherman, M., Paulauskis, J., Millis, S., Kahn, A., Greene, J. M., Sfeir, R., Jensen, M. A., Chen, J., Whitmore, J., Alonso, S., Jordan, J., Chu, A., Zhang, J., Barker, A., Compton, C., Eley, G., Ferguson, M., Fielding, P., Gerhard, D. S., Myles, R., Schaefer, C., Shaw, K. R. M., Vaught, J., Vockley, J. B., Good, P. J., Guyer, M. S., Ozenberger, B., Peterson, J. & Thomson, E. (2011). Nature 474, 609–615.

Bowtell, D. D., Böhm, S., Ahmed, A. A., Aspuria, P.-J., Bast, R. C., Beral, V., Berek, J. S., Birrer, M. J., Blagden, S., Bookman, M. A., Brenton, J. D., Chiappinelli, K. B., Martins, F. C., Coukos, G., Drapkin, R., Edmondson, R., Fotopoulou, C., Gabra, H., Galon, J., Gourley, C., Heong, V., Huntsman, D. G., Iwanicki, M., Karlan, B. Y., Kaye, A., Lengyel, E., Levine, D. A., Lu, K. H., McNeish, I. A., Menon, U., Narod, S. A., Nelson, B. H., Nephew, K. P., Pharoah, P., Powell, D. J., Ramos, P., Romero, I. L., Scott, C. L., Sood, A. K., Stronach, E. A. & Balkwill, F. R. (2015). Nature Reviews Cancer 15, 668–679.

Boyle, S. E., Fedele, C. G., Corbin, V., Wybacz, E., Szeto, P., Lewin, J., Young, R. J.,

Wong, A., Fuller, R., Spillane, J., Speakman, D., Donahoe, S., Pohl, M., Gyorki, D., Henderson, M. A., Johnstone, R. W., Papenfuss, A. T. & Shackleton, M. (2016). Cancer Research 76, 3965–3977.

Brown, J. & Frumovitz, M. (2014). Current Oncology Reports 16.

Buczacki, S. J. A., Zecchini, H. I., Nicholson, A. M., Russell, R., Vermeulen, L., Kemp, R. & Winton, D. J. (2013). Nature 495, 65–69.

Campbell, I. G., Russell, S. E., Choong, D. Y. H., Montgomery, K. G., Ciavarella, M. L., Hooi, C. S. F., Cristiano, B. E., Pearson, R. B. & Phillips, W. A. (2004). Cancer Research 64, 7678–7681.

Casagrande, J., Pike, M., Ross, P., Louie, E., Roy, S. & Henderson, B. (1979). The Lancet 314, 170–173.

Chandler, R. L., Damrauer, J. S., Raab, J. R., Schisler, J. C., Wilkerson, M. D., Didion, J. P., Starmer, J., Serber, D., Yee, D., Xiong, J., Darr, D. B., de Villena, F. P.-M., Kim, W. Y. & Magnuson, T. (2015). Nature Communications 6.

Chen, J., Li, Y., Yu, T.-S., McKay, R. M., Burns, D. K., Kernie, S. G. & Parada, L. F. (2012). Nature 488, 522–526.

Chen, M., Yao, S., Cao, Q., Xia, M., Liu, J. & He, M. (2017). Oncotarget 8.

Chene, G., Tchirkov, A., Pierre-Eymard, E., Dauplat, J., Raoelfils, I., Cayre, A., Watkin, E., Vago, P. & Penault-Llorca, F. (2013). Clinical Cancer Research 19, 2873–2882.

Clayton, E., Doupé, D. P., Klein, A. M., Winton, D. J., Simons, B. D. & Jones, P. H. (2007). Nature 446, 185–189.

Comamala, M., Pinard, M., Thériault, C., Matte, I., Albert, A., Boivin, M., Beaudin, J., Piché, A. & Rancourt, C. (2011). British Journal of Cancer 104, 989–999.

Curley, M. D., Therrien, V. A., Cummings, C. L., Sergent, P. A., Koulouris, C. R., Friel, A. M., Roberts, D. J., Seiden, M. V., Scadden, D. T., Rueda, B. R. & Foster, R. (2009). Stem cells (Dayton, Ohio) 27, 2875–2883.

Curtis, C., , Shah, S. P., Chin, S.-F., Turashvili, G., Rueda, O. M., Dunning, M. J., Speed, D., Lynch, A. G., Samarajiwa, S., Yuan, Y., Gräf, S., Ha, G., Haffari, G., Bashashati, A., Russell, R., McKinney, S., Langerød, A., Green, A., Provenzano, E., Wishart, G., Pinder, S., Watson, P., Markowetz, F., Murphy, L., Ellis, I., Purushotham, A., Børresen-Dale, A.-L., Brenton, J. D., Tavaré, S., Caldas, C. & Aparicio, S. (2012). Nature 486, 346–352.

Cybulska, P., Stewart, J. M., Sayad, A., Virtanen, C., Shaw, P. A., Clarke, B., Stickle, N., Bernardini, M. Q. & Neel, B. G. (2018). The American Journal of Pathology 188, 1120–1131.

Davy, M., Mossige, J. & Johannessen, J. V. (1977). Acta Obstetricia et Gynecologica Scandinavica 56, 55–59.

Driessens, G., Beck, B., Caauwe, A., Simons, B. D. & Blanpain, C. (2012). Nature 488, 527–530.

Dubeau, L. (1999). Gynecologic Oncology 72, 437–442.

Eckert, M. A., Pan, S., Hernandez, K. M., Loth, R. M., Andrade, J., Volchenboum, S. L., Faber, P., Montag, A., Lastra, R., Peter, M. E., Yamada, S. D. & Lengyel, E. (2016). Cancer Discovery 6, 1342–1351.

Edwards, S. L., Brough, R., Lord, C. J., Natrajan, R., Vatcheva, R., Levine, D. A., Boyd, J., Reis-Filho, J. S. & Ashworth, A. (2008). Nature 451, 1111–1115.

Eirew, P., Stingl, J., Raouf, A., Turashvili, G., Aparicio, S., Emerman, J. T. & Eaves, C. J. (2008). Nature Medicine 14, 1384–1389.

Erba, E., Giordano, M., Danova, M., Mazzini, G., Ubezio, P., Torri, V., Mangioni, C., Landoni, F., Bolis, P. & Tenti, P. (1994). Annals of oncology : official journal of the European Society for Medical Oncology 5, 627–634.

Flesken-Nikitin, A., Hwang, C.-I., Cheng, C.-Y., Michurina, T. V., Enikolopov, G. & Nikitin, A. Y. (2013). Nature 495, 241–245.

Forshew, T., Murtaza, M., Parkinson, C., Gale, D., Tsui, D. W. Y., Kaper, F., Dawson, S.-J., Piskorz, A. M., Jimenez-Linan, M., Bentley, D., Hadfield, J., May, A. P., Caldas, C., Brenton, J. D. & Rosenfeld, N. (2012). Science Translational Medicine 4, 136ra68–136ra68.

George, E., Kim, H., Krepler, C., Wenz, B., Makvandi, M., Tanyi, J. L., Brown, E., Zhang, R., Brafford, P., Jean, S., Mach, R. H., Lu, Y., Mills, G. B., Herlyn, M., Morgan, M., Zhang, X., Soslow, R., Drapkin, R., Johnson, N., Zheng, Y., Cotsarelis, G., Nathanson, K. L. & Simpkins, F. (2017). JCI Insight 2.

Gershenson, D. M. (2010). Cancer 116, 1400–1402.

Giannakouros, P., Matte, I., Rancourt, C. & piché, A. (2014). International Journal of Oncology 46, 91–98.

Gilks, C. B. (2004). International Journal of Gynecological Pathology 23, 200–205.

Giraddi, R. R., Shehata, M., Gallardo, M., Blasco, M. A., Simons, B. D. & Stingl, J. (2015). Nature Communications 6, 8487.

Hanahan, D. & Weinberg, R. A. (2000). Cell 100, 57–70.

Hidalgo, M., Amant, F., Biankin, A. V., Budinska, E., Byrne, A. T., Caldas, C., Clarke, R. B., de Jong, S., Jonkers, J., Maelandsmo, G. M., Roman-Roman, S., Seoane, J., Trusolino, L. & and, A. V. (2014). Cancer Discovery 4, 998–1013.

Hoshino, A., Ishii, G., Ito, T., Aoyagi, K., Ohtaki, Y., Nagai, K., Sasaki, H. & Ochiai, A. (2011). Cancer Research 71, 4769–4779.

Huynh, P. T., Beswick, E. J., Coronado, Y. A., Johnson, P., O'Connell, M. R., Watts, T., Singh, P., Qiu, S., Morris, K., Powell, D. W. & Pinchuk, I. V. (2015). International Journal of Cancer 138, 1971–1981.

Indraccolo, S., Habeler, W., Tisato, V., Stievano, L., Piovan, E., Tosello, V., Esposito, G., Wagner, R., Uberla, K., Chieco-Bianchi, L. & Amadori, A. (2002). Cancer research 62, 6099–6107.

Janzen, D. M., Tiourin, E., Salehi, J. A., Paik, D. Y., Lu, J., Pellegrini, M. & Memarzadeh, S. (2015). Nature Communications 6, 7956.

Janzen, D. M., Tiourin, E., Salehi, J. A., Paik, D. Y., Lu, J., Pellegrini, M. & Memarzadeh, S. (2016). Nature Communications 7, 10703.

Jones, S., Wang, T.-L., Shih, I.-M., Mao, T.-L., Nakayama, K., Roden, R., Glas, R., Slamon, D., Diaz, L. A., Vogelstein, B., Kinzler, K. W., Velculescu, V. E. & Papadopoulos, N. (2010). Science 330, 228–231.

Kaldawy, A., Segev, Y., Lavie, O., Auslender, R., Sopik, V. & Narod, S. A. (2016). Gynecologic Oncology 143, 433–438.

Kamatchi, V., Babu, N., Sankari, S. & Rajesh, E. (2015). Journal of Pharmacy and Bioallied Sciences 7, 209.

Kawase, A., Ishii, G., Nagai, K., Ito, T., Nagano, T., Murata, Y., Hishida, T., Nishimura, M., Yoshida, J., Suzuki, K. & Ochiai, A. (2008). International Journal of Cancer 123, 1053–1059.

Kerr, J. F. R., Wyllie, A. H. & Currie, A. R. (1972). British Journal of Cancer 26, 239–257.

Konecny, G. E. & Kristeleit, R. S. (2016). British Journal of Cancer 115, 1157–1173.

Kozar, S., Morrissey, E., Nicholson, A. M., van der Heijden, M., Zecchini, H. I., Kemp, R., Tavaré, S., Vermeulen, L. & Winton, D. J. (2013). Cell Stem Cell 13, 626–633.

Kryczek, I., Liu, S., Roh, M., Vatan, L., Szeliga, W., Wei, S., Banerjee, M., Mao, Y., Kotarski, J., Wicha, M. S., Liu, R. & Zou, W. (2011). International Journal of Cancer 130, 29–39.

Kurman, R. J. & Shih, I.-M. (2010). The American Journal of Surgical Pathology 34, 433–443.

Kurman, R. J. & Shih, I.-M. (2011). Human Pathology 42, 918–931.

Kyritsis, A. P., Bondy, M. L., Hess, K. R., Cunningham, J. E., Zhu, D., Amos, C. J., Yung, W. K., Levin, V. A. & Bruner, J. M. (1995). Clinical cancer research : an official journal of the American Association for Cancer Research 1, 1617–1622.

Landen, C. N., Birrer, M. J. & Sood, A. K. (2008). Journal of Clinical Oncology 26, 995–1005.

Lawson, D. A., Bhakta, N. R., Kessenbrock, K., Prummel, K. D., Yu, Y., Takai, K., Zhou, A., Eyob, H., Balakrishnan, S., Wang, C.-Y., Yaswen, P., Goga, A. & Werb, Z. (2015). Nature 526, 131–135.

Lee, Y., Miron, A., Drapkin, R., Nucci, M., Medeiros, F., Saleemuddin, A., Garber, J., Birch, C., Mou, H., Gordon, R., Cramer, D., McKeon, F. & Crum, C. (2006). The Journal of Pathology 211, 26–35.

Liu, A. Y., Roudier, M. P. & True, L. D. (2004). The American Journal of Pathology 165, 1543–1556.

Liu, J. F., Palakurthi, S., Zeng, Q., Zhou, S., Ivanova, E., Huang, W., Zervantonakis, I. K., Selfors, L. M., Shen, Y., Pritchard, C. C., Zheng, M., Adleff, V., Papp, E., Piao, H., Novak, M., Fotheringham, S., Wulf, G. M., English, J., Kirschmeier, P. T., Velculescu, V. E., Paweletz, C., Mills, G. B., Livingston, D. M., Brugge, J. S., Matulonis, U. A. & Drapkin, R. (2016). Clinical Cancer Research 23, 1263–1273.

Macintyre, G., Goranova, T. E., Silva, D. D., Ennis, D., Piskorz, A. M., Eldridge, M., Sie, D., Lewsley, L.-A., Hanif, A., Wilson, C., Dowson, S., Glasspool, R. M., Lockley, M., Brockbank, E., Montes, A., Walther, A., Sundar, S., Edmondson, R., Hall, G. D., Clamp, A., Gourley, C., Hall, M., Fotopoulou, C., Gabra, H., Paul, J., Supernat, A., Millan, D., Hoyle, A., Bryson, G., Nourse, C., Mincarelli, L., Sanchez, L. N., Ylstra, B., Jimenez-Linan, M., Moore, L., Hofmann, O., Markowetz, F., McNeish, I. A. & Brenton, J. D. (2018). Nature Genetics 50, 1262–1270.

Mao, T.-L. & Shih, I.-M. (2013). Journal of Gynecologic Oncology 24, 376.

McGuire, V., Jesser, C. A. & Whittemore, A. S. (2002). Gynecologic Oncology 84, 399–403.

Medeiros, F., Muto, M. G., Lee, Y., Elvin, J. A., Callahan, M. J., Feltmate, C., Garber, J. E., Cramer, D. W. & Crum, C. P. (2006). The American Journal of Surgical Pathology 30, 230–236.

Mittempergher, L. (2016). Current Oncology Reports 18.

Moitra, K. (2015). BioMed Research International 2015, 1–8.

Morton, C. L. & Houghton, P. J. (2007). Nature Protocols 2, 247–250.

Paik, D. Y., Janzen, D. M., Schafenacker, A. M., Velasco, V. S., Shung, M. S., Cheng, D., Huang, J., Witte, O. N. & Memarzadeh, S. (2012). STEM CELLS 30, 2487–2497.

Pastushenko, I., Brisebarre, A., Sifrim, A., Fioramonti, M., Revenco, T., Boumahdi, S., Keymeulen, A. V., Brown, D., Moers, V., Lemaire, S., Clercq, S. D., Minguijón, E., Balsat, C., Sokolow, Y., Dubois, C., Cock, F. D., Scozzaro, S., Sopena, F., Lanas, A., D'Haene, N., Salmon, I., Marine, J.-C., Voet, T., Sotiropoulou, P. A. & Blanpain, C. (2018). Nature 556, 463–468.

Perets, R., Wyant, G. A., Muto, K. W., Bijron, J. G., Poole, B. B., Chin, K. T., Chen, J. Y. H., Ohman, A. W., Stepule, C. D., Kwak, S., Karst, A. M., Hirsch, M. S., Setlur, S. R., Crum, C. P., Dinulescu, D. M. & Drapkin, R. (2013). Cancer Cell 24, 751–765.

Piskorz, A. M., Ennis, D., Macintyre, G., Goranova, T. E., Eldridge, M., Segui-Gracia, N., Valganon, M., Hoyle, A., Orange, C., Moore, L., Jimenez-Linan, M., Millan, D., McNeish, I. A. & Brenton, J. D. (2015). Annals of Oncology 27, 532–539.

Potten, C. (1996). British Journal of Cancer 74, 1743–1748.

Potten, C. S. (1977). Nature 269, 518–521.

Potten, C. S. (1992). Cancer and Metastasis Review 11, 179–195.

Potten, C. S., Al-Barwari, S. E., Hume, W. J. & Searle, J. (1977). Cell Proliferation 10, 557–568.

Qiao, L., Koutsos, M., Tsai, L.-L., Kozoni, V., Guzman, J., Shiff, S. J. & Rigas, B. (1996). Cancer Letters 107, 83–89.

Quintana, E., Shackleton, M., Sabel, M. S., Fullen, D. R., Johnson, T. M. & Morrison, S. J. (2008). Nature 456, 593–598.

Reed, S. I., Spruck, C. H. & Won, K.-A. (1999). Nature 401, 297–300.

Rehm, M., Düßmann, H., Jänicke, R. U., Tavaré, J. M., Kögel, D. & Prehn, J. H. M. (2002). Journal of Biological Chemistry 277, 24506–24514.

Rew, D. & Wilson, G. (2000). European Journal of Surgical Oncology (EJSO) 26, 405–417.

Rios, A. C., Fu, N. Y., Lindeman, G. J. & Visvader, J. E. (2014). Nature 506, 322–327.

Rodriguez, E., Mannion, L., D'Santos, P., Griffiths, M., Arends, M. J., Brindle, K. M. & Lyons, S. K. (2014). The Journal of Pathology 232, 449–457.

Rosen, D. G., Yang, G., Deavers, M. T., Malpica, A., Kavanagh, J. J., Mills, G. B. & Liu, J. (2006). Cancer 106, 1925–1932.

Sallum, L. F., Andrade, L., da Costa, L. B. E., Ramalho, S., Ferracini, A. C., de Andrade Natal, R., Brito, A. B. C., Sarian, L. O. & Derchain, S. (2018). International Journal of Gynecological Cancer 28, 437–447.

Schoppmann, S. F., Berghoff, A., Dinhof, C., Jakesz, R., Gnant, M., Dubsky, P., Jesch, B., Heinzl, H. & Birner, P. (2012). Breast Cancer Research and Treatment 134, 237–244.

Shayesteh, L., Lu, Y., Kuo, W.-L., Baldocchi, R., Godfrey, T., Collins, C., Pinkel, D., Powell, B., Mills, G. B. & Gray, J. W. (1999). Nature Genetics 21, 99–102.

Sieh, W., Köbel, M., Longacre, T. A., Bowtell, D. D., deFazio, A., Goodman, M. T., Høgdall, E., Deen, S., Wentzensen, N., Moysich, K. B., Brenton, J. D., Clarke, B. A., Menon, U., Gilks, C. B., Kim, A., Madore, J., Fereday, S., George, J., Galletta, L., Lurie, G., Wilkens, L. R., Carney, M. E., Thompson, P. J., Matsuno, R. K., Kjær, S. K., Jensen, A., Høgdall, C., Kalli, K. R., Fridley, B. L., Keeney, G. L., Vierkant, R. A., Cunningham, J. M., Brinton, L. A., Yang, H. P., Sherman, M. E., García-Closas, M., Lissowska, J., Odunsi, K., Morrison, C., Lele, S., Bshara, W., Sucheston, L., Jimenez-Linan, M., Driver, K., Alsop, J., Mack, M., McGuire, V., Rothstein, J. H., Rosen, B. P., Bernardini, M. Q., Mackay, H., Oza, A., Wozniak, E. L., Benjamin, E., Gentry-Maharaj, A., Gayther, S. A., Tinker, A. V., Prentice, L. M., Chow, C., Anglesio, M. S., Johnatty, S. E., Chenevix-Trench, G., Whittemore, A. S., Pharoah, P. D., Goode, E. L., Huntsman, D. G. & Ramus, S. J. (2013). The Lancet Oncology 14, 853–862.

Silva, I. A., Bai, S., McLean, K., Yang, K., Griffith, K., Thomas, D., Ginestier, C., Johnston, C., Kueck, A., Reynolds, R. K., Wicha, M. S. & Buckanovich, R. J. (2011). Cancer Research 71, 3991–4001.

Soslow, R. A. (2008). International Journal of Gynecological Pathology PAP.

Stewart, J. M., Shaw, P. A., Gedye, C., Bernardini, M. Q., Neel, B. G. & Ailles, L. E. (2011). Proceedings of the National Academy of Sciences 108, 6468–6473.

Szotek, P. P., Pieretti-Vanmarcke, R., Masiakos, P. T., Dinulescu, D. M., Connolly, D., Foster, R., Dombkowski, D., Preffer, F., MacLaughlin, D. T. & Donahoe, P. K. (2006). Proceedings of the National Academy of Sciences 103, 11154–11159.

Tawa, P., Hell, K., Giroux, A., Grimm, E., Han, Y., Nicholson, D. W. & Xanthoudakis, S. (2004). Cell Death and Differentiation 11, 439–447.

Thornberry, N. A. (1998). Science 281, 1312–1316.

Van Keymeulen, A., Rocha, A. S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S. & Blanpain, C. (2011). Nature 479, 189–193.

Vaughan, S., Coward, J. I., Bast, R. C., Berchuck, A., Berek, J. S., Brenton, J. D., Coukos, G., Crum, C. C., Drapkin, R., Etemadmoghadam, D., Friedlander, M., Gabra, H., Kaye, S. B., Lord, C. J., Lengyel, E., Levine, D. A., McNeish, I. A., Menon, U., Mills, G. B., Nephew, K. P., Oza, A. M., Sood, A. K., Stronach, E. A., Walczak, H., Bowtell, D. D. & Balkwill, F. R. (2011). Nature Reviews Cancer 11, 719–725.

Vermeulen, L., Morrissey, E., van der Heijden, M., Nicholson, A. M., Sottoriva, A., Buczacki, S., Kemp, R., Tavare, S. & Winton, D. J. (2013). Science 342, 995–998.

Wei, X., Dombkowski, D., Meirelles, K., Pieretti-Vanmarcke, R., Szotek, P. P., Chang, H. L., Preffer, F. I., Mueller, P. R., Teixeira, J., MacLaughlin, D. T. & Donahoe, P. K. (2010). Proceedings of the National Academy of Sciences 107, 18874–18879.

Weroha, S. J., Becker, M. A., Enderica-Gonzalez, S., Harrington, S. C., Oberg, A. L., Maurer, M. J., Perkins, S. E., AlHilli, M., Butler, K. A., McKinstry, S., Fink, S., Jenkins, R. B., Hou, X., Kalli, K. R., Goodman, K. M., Sarkaria, J. N., Karlan, B. Y., Kumar, A., Kaufmann, S. H., Hartmann, L. C. & Haluska, P. (2014). Clinical Cancer Research 20, 1288–1297.

Wiegand, K. C., Shah, S. P., Al-Agha, O. M., Zhao, Y., Tse, K., Zeng, T., Senz, J., McConechy, M. K., Anglesio, M. S., Kalloger, S. E., Yang, W., Heravi-Moussavi, A., Giuliany, R., Chow, C., Fee, J., Zayed, A., Prentice, L., Melnyk, N., Turashvili, G., Delaney, A. D., Madore, J., Yip, S., McPherson, A. W., Ha, G., Bell, L., Fereday, S., Tam, A., Galletta, L., Tonin, P. N., Provencher, D., Miller, D., Jones, S. J., Moore, R. A., Morin, G. B., Oloumi, A., Boyd, N., Aparicio, S. A., Shih, I.-M., Mes-Masson, A.-M., Bowtell, D. D., Hirst, M., Gilks, B., Marra, M. A. & Huntsman, D. G. (2010). New England Journal of Medicine 363, 1532–1543.

Winterhoff, B. J., Maile, M., Mitra, A. K., Sebe, A., Bazzaro, M., Geller, M. A., Abrahante, J. E., Klein, M., Hellweg, R., Mullany, S. A., Beckman, K., Daniel, J. & Starr, T. K. (2017). Gynecologic oncology 144, 598–606.

Yamanashi, T., Nakanishi, Y., Fujii, G., Akishima-Fukasawa, Y., Moriya, Y., Kanai, Y., Watanabe, M. & Hirohashi, S. (2009). Oncology 77, 53–62.

Yang-Hartwich, Y., Soteras, M. G., Lin, Z. P., Holmberg, J., Sumi, N., Craveiro, V., Liang, M., Romanoff, E., Bingham, J., Garofalo, F., Alvero, A. & Mor, G. (2014). Oncogene 34, 3605–3616.

Zhang, S., Balch, C., Chan, M. W., Lai, H.-C., Matei, D., Schilder, J. M., Yan, P. S., Huang, T. H.-M. & Nephew, K. P. (2008). Cancer Research 68, 4311–4320.

Zhu, Y., Huang, Y.-F., Kek, C. & Bulavin, D. V. (2013). Cell Stem Cell 12, 298–303.